

79280

184707

Mg

STIC-Biotech/ChemLib

From: Swope, Sheridan
Sent: Friday, April 07, 2006 7:23 PM
To: STIC-Biotech/ChemLib
Subject: 10/825,632

For 10/825,632 , pls search and interference search:

SID 1 against the NT and AA databases.

SID 2 against the NT and AA databases.

Thanks!

Sheridan Swope, Ph.D.
Primary Patent Examiner
AU 1656/Recombinant Enzymes
571-272-0943 (voice)
E02B71 Remsen Bld (Office)
E03C70 Remsen Bld (Mailbox)

Not 4/27/06
6/20/06

RECEIVED
APR 10 2006
STIC

BEST AVAILABLE COPY

Searcher: _____
Searcher Phone: _____
Date Searcher Picked up: _____
Date completed: _____
Searcher Prep Time: _____
Online Time: _____

Type of Search
NA# _____ AA# _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure #: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable
STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other (Specify): _____

THIS PAGE BLANK (USPTO)



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 184707

TO: Sheridan Swope
Location: rem/2B71/3C70
Art Unit: 1656
Thursday, April 20, 2006
Case Serial Number: 10/825632

From: Kristine Hensle
Location: Biotech-Chem Library
REM-1B69
Phone: (571) 272-4161

Kristine. Hensle@uspto.gov

Search Notes

Examiner Swope,

See attached results.

If you have any questions about this search feel free to contact me at any time.

Thank you for using STIC search services!

Kristine Hensle
Librarian
STIC Biotech/Chem Library
(571)272-4161



THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 14, 2006, 12:42:35 ; Search time 193 Seconds
(without alignments)
2007.938 Million cell updates/sec

Title: US-10-825-632-1

Perfect score: 4700

Sequence: 1 MAAMETEQLGVETATDC.....HLLHYLQENLGSRIALKVI 882

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A Geneseq 21.*

1: Geneseqp1980s.*
2: Geneseqp1990s.*
3: Geneseqp2000s.*
4: Geneseqp2001s.*
5: Geneseqp2002s.*
6: Geneseqp2003as.*
7: Geneseqp2003bs.*
8: Geneseqp2004s.*
9: Geneseqp2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	4700	100.0	882	4 AAB47187	AAB47187 Human DPP
2	4700	100.0	882	5 ABG61591	ABG61591 Human DPP
3	4700	100.0	882	5 AAG78415	AAG78415 Amino aci
4	4700	100.0	882	5 AAE24170	AAE24170 Human dip
5	4700	100.0	882	5 AAU74749	AAU74749 Human pro
6	4700	100.0	882	5 ADI17086	ADI17086 Human NOV
7	4700	100.0	882	6 ABU07720	ABU07720 Human ser
8	4528.5	96.4	883	5 ADI17085	ADI17085 Murine NO
9	4397.5	93.6	831	6 ABU92026	ABU92026 Human pro
10	4353	92.6	824	6 ABU92030	ABU92030 Human pro
11	4118	87.6	782	5 ABG97361	ABG97361 Novel hum
12	3901	83.0	746	8 ADI16334	ADI16334 Human pro
13	3817	81.2	738	8 ADT04044	ADT04044 Human pro
14	3771	80.2	724	5 ABG97362	ABG97362 Novel hum
15	3504	74.6	658	5 ABG61600	ABG61600 Human DPP
16	3504	74.6	661	5 ABG61596	ABG61596 Human DPP
17	3504	74.6	690	5 ABG61594	ABG61594 Human DPP
18	3333.5	70.9	632	4 AAB93565	AAB93565 Human pro
19	3236	68.9	613	5 ABG61601	ABG61601 Human DPP
20	3087	65.7	587	5 ADRA1398	ADRA1398 Human CD-
21	2870	61.1	863	5 ABG61592	ABG61592 Human DPP
22	2870	61.1	863	5 ADI17083	ADI17083 Human NOV
23	2870	61.1	892	5 ABG61602	ABG61602 Human DPP
24	2870	61.1	892	5 ABG61604	ABG61604 Human DPP

25	2870	61.1	892	5 ABB98134	Abb98134 Human PMW
26	2870	61.1	892	8 ADS10951	Ads10951 Human the
27	2863	60.9	969	5 AAE24168	Aae24168 Human dip
28	2862	60.9	863	5 ADI16690	Adi16690 Human NOV
29	2862	60.9	863	5 ADI16688	Adi16688 Human NOV
30	2862	60.9	863	8 ADN42344	Adn42344 Human nov
31	2862	60.9	863	8 ADN42342	Adn42342 Human nov
32	2835	60.3	830	5 AAE24171	Aae24171 Human dip
33	2833	60.3	869	5 AAE24169	Aae24169 Alternati
34	2820.5	60.0	879	5 ABG61607	ABG61607 Human DPP
35	2820.5	60.0	879	5 ABG61608	ABG61608 Human DPP
36	2820.5	60.0	879	8 ADS10953	Ads10953 Human the
37	2713.5	57.7	847	5 AAE23875	Aae23875 Murine di
38	2707	57.6	854	5 ADS10952	Ads10952 Human the
39	2671.5	56.8	873	8 ADS10479	Ads10479 Human the
40	2547.5	54.2	580	5 AAE14337	Aae14337 Human pro
41	2498	53.1	755	8 ADO67811	Ado67811 Novel hum
42	2422	51.5	465	4 AAB47189	Aab47189 Human DPP
43	2413.5	51.4	737	4 AAM38724	Aam38724 Human pol
44	2406	51.2	832	5 ABG61605	ABG61605 Human DPP
45	2406	51.2	832	5 ABG61606	ABG61606 Human DPP

ALIGNMENTS

RESULT 1

AAB47187

ID AAB47187 standard; protein; 882 AA.

XX AC AAB47187;

XX AC AAB47187;

DT 29-JUN-2001 (first entry)

XX DE Human DPP8.

XX DE Human DPP8.

XX KW Human; dipeptidyl aminopeptidase; DPP8; prolol oligopeptidase;

XX KW dipeptidyl peptidase; DPPIV; T cell; cleavage; diarrhoea;

XX KW growth hormone deficiency; glucose level; mucosal regeneration;

XX KW non-insulin dependent diabetes mellitus; glucose intolerance;

XX KW immunosuppression;

XX OS Homo sapiens.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX FT Active-site /note= "Forms part of Ser-Asp-His catalytic triad"

XX FT Active-site /note= "Forms part of Ser-Asp-His catalytic triad"

XX FT Active-site /note= "Forms part of Ser-Asp-His catalytic triad"

XX FT Active-site /note= "Forms part of Ser-Asp-His catalytic triad"

XX FT Active-site /note= "Forms part of Ser-Asp-His catalytic triad"

XX FT Active-site /note= "Forms part of Ser-Asp-His catalytic triad"

XX FT Active-site /note= "Forms part of Ser-Asp-His catalytic triad"

XX FT Active-site /note= "Forms part of Ser-Asp-His catalytic triad"

XX FT Active-site /note= "Forms part of Ser-Asp-His catalytic triad"

XX FT Active-site /note= "Forms part of Ser-Asp-His catalytic triad"

XX FT Active-site /note= "Forms part of Ser-Asp-His catalytic triad"

XX FT Active-site /note= "Forms part of Ser-Asp-His catalytic triad"

XX FT Active-site /note= "Forms part of Ser-Asp-His catalytic triad"

XX FT Active-site /note= "Forms part of Ser-Asp-His catalytic triad"

XX FT Active-site /note= "Forms part of Ser-Asp-His catalytic triad"

XX FT Active-site /note= "Forms part of Ser-Asp-His catalytic triad"

XX FT Active-site /note= "Forms part of Ser-Asp-His catalytic triad"

XX FT Active-site /note= "Forms part of Ser-Asp-His catalytic triad"

XX FT Active-site /note= "Forms part of Ser-Asp-His catalytic triad"

XX FT Active-site /note= "Forms part of Ser-Asp-His catalytic triad"

XX FT Active-site /note= "Forms part of Ser-Asp-His catalytic triad"

XX FT Active-site /note= "Forms part of Ser-Asp-His catalytic triad"

XX FT Active-site /note= "Forms part of Ser-Asp-His catalytic triad"

XX FT Active-site /note= "Forms part of Ser-Asp-His catalytic triad"

XX FT Active-site /note= "Forms part of Ser-Asp-His catalytic triad"

XX FT Active-site /note= "Forms part of Ser-Asp-His catalytic triad"

XX FT Active-site /note= "Forms part of Ser-Asp-His catalytic triad"

XX FT Active-site /note= "Forms part of Ser-Asp-His catalytic triad"

XX FT Active-site /note= "Forms part of Ser-Asp-His catalytic triad"

XX FT Active-site /note= "Forms part of Ser-Asp-His catalytic triad"

XX FT Active-site /note= "Forms part of Ser-Asp-His catalytic triad"

XX FT Active-site /note= "Forms part of Ser-Asp-His catalytic triad"

XX FT Active-site /note= "Forms part of Ser-Asp-His catalytic triad"

XX FT Active-site /note= "Forms part of Ser-Asp-His catalytic triad"

Db 1 MAAAMETEQLGVEIFETADCEENIESQDRPKLEPFYVERYSWSQKLLADTRKYHYMM 60
Qy 61 AKAPHDFMFKVNDPDGPHSDRIYYLAMSNGENRENTLYFSEIYKTNRAAVLMLSKPPL 120
Db 61 AKAPHDFMFKVNDPDGPHSDRIYYLAMSNGENRENTLYFSEIYKTNRAAVLMLSKPPL 120
Qy 121 DLFOATLDYGMYSREELLRERKRIGTVGIASDYHOGSGTFLFOAGSGIYHVKGDPG 180
Db 121 DLFOATLDYGMYSREELLRERKRIGTVGIASDYHOGSGTFLFOAGSGIYHVKGDPG 180
Qy 181 FTQOPLRPNLVETSCPNIRMDPKLCPADPDWIAFHNSDIWISNIVTREERLTYVHNE 240
Db 181 FTQOPLRPNLVETSCPNIRMDPKLCPADPDWIAFHNSDIWISNIVTREERLTYVHNE 240
Qy 241 ANNEEDARSAGVATVLOEEDRYSGYWCPCAKTTPSGGKILRIIYEENDESEVEIHV 300
Db 241 ANNEEDARSAGVATVLOEEDRYSGYWCPCAKTTPSGGKILRIIYEENDESEVEIHV 300
Qy 301 TSPMLETRADSFYKPTGTANPKVTFKQSEIMIDAGRIIDVIDKELIQPPEILPEGVE 360
Db 301 TSPMLETRADSFYKPTGTANPKVTFKQSEIMIDAGRIIDVIDKELIQPPEILPEGVE 360
Qy 361 YIARAGWPEGKYANSIILDSOTELQIVLISPELFTPVEDDWMERQRLIESVPDSVTP 420
Db 361 YIARAGWPEGKYANSIILDSOTELQIVLISPELFTPVEDDWMERQRLIESVPDSVTP 420
Qy 421 IYEEETDIIWNIHDIHVFPQSHHEEIEFIPASECKTGPRHLKYKITSILKESKYRSG 480
Db 421 IYEEETDIIWNIHDIHVFPQSHHEEIEFIPASECKTGPRHLKYKITSILKESKYRSG 480
Qy 481 GLPAPSDPKPIKEBIAITSGEWEVLGRHGSNIQVDEVRRLVYFEGTKDSPLEHLYVVS 540
Db 481 GLPAPSDPKPIKEBIAITSGEWEVLGRHGSNIQVDEVRRLVYFEGTKDSPLEHLYVVS 540
Qy 541 YNPGEVRLTDGRYSHSCCISORCDPFIISKYSNQKNPHCVSLYKLSPPEDDPTCKTEF 600
Db 541 YNPGEVRLTDGRYSHSCCISORCDPFIISKYSNQKNPHCVSLYKLSPPEDDPTCKTEF 600
Qy 601 WATILDSAGLPDPYTPPEIFSFESTGTFTLYGMLYKPHDLQPKKYPTVLFIYGGPQVL 660
Db 601 WATILDSAGLPDPYTPPEIFSFESTGTFTLYGMLYKPHDLQPKKYPTVLFIYGGPQVL 660
Qy 661 VNNRFGVKYFRLNTLASIGYVVVINDNRGSHRGLKFEAGFYKMGQIEIDDDQVEGLQY 720
Db 661 VNNRFGVKYFRLNTLASIGYVVVINDNRGSHRGLKFEAGFYKMGQIEIDDDQVEGLQY 720
Qy 721 LASRYDFIDLDRVGIHGSYGYLSLMLMORSQIFRVAIAGAPVTLWIFDTGYTERYM 780
Db 721 LASRYDFIDLDRVGIHGSYGYLSLMLMORSQIFRVAIAGAPVTLWIFDTGYTERYM 780
Qy 781 GHPDQEQGYLGSVAMQAEKFPSEPNRLLLHGFLENVHFAHTSILLSLFLVRAGKPYD 840
Db 781 GHPDQEQGYLGSVAMQAEKFPSEPNRLLLHGFLENVHFAHTSILLSLFLVRAGKPYD 840
Qy 841 LQIYQERHSIRVPSGSHYELHLHLVYLQENLGSRIAALKVI 882
Db 841 LQIYQERHSIRVPSGSHYELHLHLVYLQENLGSRIAALKVI 882

RESULT 3

AAAG78415
ID AAG78415 standard; protein; 882 AA.
XX
AC AAG78415;
XX
DT 12-APR-2002 (first entry)
XX
DE Amino acid sequence of 21953 human prolyl oligopeptidase.
XX
KW 21953 prolyl oligopeptidase; antibody; proline; endopeptidase; cancer;
KW cardiovascular disease; autoimmune disease; atopic allergy;
KW neuronal disorder; vascular disorder; prostate disorder; cytostatic;

KW antidiabetic; antiarthritic; antiasthmatic; antiinflammatory;
KW diabetes mellitus; arthritis; multiple sclerosis; asthma;
KW Grave's disease; neuronal disorder; demyelinating disease.
XX Homo sapiens.
XX WO200179473-A2.
XX 25-OCT-2001.
XX 11-APR-2001; 2001WO-US040483.
XX 18-APR-2000; 2000US-0197508P.
XX (MILL-) MILLENNIUM PHARM INC.
XX Meyers RA, Williamson M;
XX WPI; 2002-034353/04.
XX N-PSDB; AAH99934.
XX New polypeptides 21953, member of human prolyl oligopeptidase family,
XX useful as diagnostic targets and therapeutic agents for controlling
XX cancer, lymphoma and leukemia.
XX Claim 1; Page 102-103; 121pp; English.
XX This invention relates to an isolated 21953 human prolyl oligopeptidase.
XX Which is cystostatic, antidiabetic, antiarthritic, neuroprotective,
XX antithyroid, dermatological, antipsoriatic, antiasthmatic,
XX ophthalmological, antiinflammatory, nootropic, antiparkinsonian,
XX anticonvulsant, gynaecological, vasotropic, antianginal, cardiac,
XX antihypertensive, anorectic and metabolic in its action. Uses include
XX gene therapy, expression or activity of 21953 protein modulator, it is
XX useful for identifying a compound which binds to it and can be used in
XX preventing, treating or detecting a cellular proliferative or
XX differentiative disorder. The 21953 molecules can act as novel diagnostic
XX targets and therapeutic agents for controlling disorders associated with
XX the aberrant activity or degradation of peptide hormones e.g., disorders
XX associated with cell differentiation and proliferation such as cancer,
XX immune function, reproductive, neurological and cardiovascular function.
XX The 21953 molecules are thus useful for treating and preventing cellular
XX proliferative and differentiative disorders, haematopoietic neoplastic
XX disorders, immune disorders such as autoimmune diseases, diabetes
XX mellitus, arthritis, multiple sclerosis, asthma, Grave's disease,
XX neuronal disorders, demyelinating diseases, vascular disorders and
XX metabolism or pain disorders. This sequence represents the amino acid
XX sequence of 21953 human prolyl oligopeptidase
XX Sequence 882 AA;

Query Match 100.0%; Score 4700; DB 5; Length 882;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 882; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MAAAMETEQLGVEIFETADCEENIESQDRPKLEPFYVERYSWSQKLLADTRKYHYMM 60
Db 1 MAAAMETEQLGVEIFETADCEENIESQDRPKLEPFYVERYSWSQKLLADTRKYHYMM 60
Qy 61 AKAPHDFMFKVNDPDGPHSDRIYYLAMSNGENRENTLYFSEIYKTNRAAVLMLSKPPL 120
Db 61 AKAPHDFMFKVNDPDGPHSDRIYYLAMSNGENRENTLYFSEIYKTNRAAVLMLSKPPL 120
Qy 121 DLFOATLDYGMYSREELLRERKRIGTVGIASDYHOGSGTFLFOAGSGIYHVKGDPG 180
Db 121 DLFOATLDYGMYSREELLRERKRIGTVGIASDYHOGSGTFLFOAGSGIYHVKGDPG 180
Qy 181 FTQOPLRPNLVETSCPNIRMDPKLCPADPDWIAFHNSDIWISNIVTREERLTYVHNE 240
Db 181 FTQOPLRPNLVETSCPNIRMDPKLCPADPDWIAFHNSDIWISNIVTREERLTYVHNE 240
Qy 241 ANNEEDARSAGVATVLOEEDRYSGYWCPCAKTTPSGGKILRIIYEENDESEVEIHV 300
XX

Db	241	ANNEEDARSAGVATFVLQEEFDYSGYWMCPKAETTPSGGKILRLIYEENDESVEIHHV	300
Qy	301	TSPMLETRRADSFYRPTGTANPKVTFKQSEIMIDAEGRIIDVIDKELIQPFELFEGVE	360
Db	301	TSPMLETRRADSFYRPTGTANPKVTFKQSEIMIDAEGRIIDVIDKELIQPFELFEGVE	360
Qy	361	YIARAGWTPPEKGYAWSILLDRSQTRLQIVLISPELFPVEDDVMERQRLIESVPDSVTPL	420
Db	361	YIARAGWTPPEKGYAWSILLDRSQTRLQIVLISPELFPVEDDVMERQRLIESVPDSVTPL	420
Qy	421	IYEETTDIWINIHDIHVFPQSHHEEIEFIFASECKTGFRHLYKITSLKESKYGRSSG	480
Db	421	IYEETTDIWINIHDIHVFPQSHHEEIEFIFASECKTGFRHLYKITSLKESKYGRSSG	480
Qy	481	GLPAPSDPKCPDKKEIAITSGEWEVLGRHGSNIQVDEVRRLVYFEGTKDSPLEHLLVYVS	540
Db	481	GLPAPSDPKCPDKKEIAITSGEWEVLGRHGSNIQVDEVRRLVYFEGTKDSPLEHLLVYVS	540
Qy	541	YVNPGEVTRLDRGYSHSCCISQHCDFFIKSYNQKNPHCVSLYKLSPPEDDPTCKTKEF	600
Db	541	YVNPGEVTRLDRGYSHSCCISQHCDFFIKSYNQKNPHCVSLYKLSPPEDDPTCKTKEF	600
Qy	601	WATILDSAGPLPDYTPPEIFSFESTTGFTLYGMLYKPHDLQPGKYPVTVLFYGGPQVQL	660
Db	601	WATILDSAGPLPDYTPPEIFSFESTTGFTLYGMLYKPHDLQPGKYPVTVLFYGGPQVQL	660
Qy	661	VNNRFGVKYFRNLNTLASGLVYVVVVVDNRGSGCHRGKLFEGAPKYKMGQIEIDDOVEGLQY	720
Db	661	VNNRFGVKYFRNLNTLASGLVYVVVVVDNRGSGCHRGKLFEGAPKYKMGQIEIDDOVEGLQY	720
Qy	721	LASRYDFIDLDRVGIHGSYGGYLSLMALMQRSDIFRVAIAGAPVTLWIFDGTGYTERYM	780
Db	721	LASRYDFIDLDRVGIHGSYGGYLSLMALMQRSDIFRVAIAGAPVTLWIFDGTGYTERYM	780
Qy	781	GHPDNEQGYLYGSAVMAQAEFPSEPNRLLHLLHVLQENLGSRIAALKVI	840
Db	781	GHPDNEQGYLYGSAVMAQAEFPSEPNRLLHLLHVLQENLGSRIAALKVI	840
Qy	841	LQIYQERHSIRVPSGHEHLLHLLHVLQENLGSRIAALKVI	882
Db	841	LQIYQERHSIRVPSGHEHLLHLLHVLQENLGSRIAALKVI	882
RESULT 4			
AAE24170			
ID	AAE24170 standard; protein; 882 AA.		
AC	AAE24170;		
XX			
DT	23-SBP-2002 (first entry)		
DE	Human dipeptidyl peptidase 8 (DPP8) protein.		
XX			
KW	Human; dipeptidyl peptidase; DPP; neoplasia; type II diabetes; cirrhosis;		
KW	autoimmunity; human immuno deficiency virus; HIV infection; cytostatic;		
KW	graft rejection; antidiabetic; antiinflammatory; immunosuppressive;		
KW	antiviral; enzyme.		
XX			
OS	Homo sapiens.		
XX			
PN	W0200234900-A1.		
XX			
PD	02-MAY-2002.		
XX			
PF	29-OCT-2001; 2001WO-AU0001388.		
XX			
PR	27-OCT-2000; 2000AU-00001078.		
XX			
PA	(UNSY) UNIV SYDNEY.		
XX			
PI	Abbott CA, Gorrell WD;		
XX			
WPI	2002-454646/48.		
DR			

DR	N-PSDB; AAD38956.			
XX	New dipeptidyl peptidase (DPP) peptidase, useful for screening inhibitors			
PT	of DPP catalytic activity, which may be employed to treat e.g. neoplasia,			
PT	type II diabetes, cirrhosis, autoimmunity, graft rejection and HIV			
PT	infection.			
XX	Example; Fig 1; 91pp; English.			
XX	The present invention relates to dipeptidyl peptidase (DPP) proteins and			
CC	polynucleotides encoding such proteins. The DPP peptides are useful for			
CC	screening inhibitors of DPP catalytic activity. The inhibitors are useful			
CC	for treating neoplasia, type II diabetes, cirrhosis, autoimmunity, graft			
CC	rejection and HIV (human immuno deficiency virus) infection. The present			
CC	sequence is human DPP8 protein			
XX	Sequence 882 AA;			
SQ	Query Match 100.0%; Score 4700; DB 5; Length 882;			
	Best Local Similarity 100.0%; Pred. No. 0;			
	Matches 882; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	MAAAMETEQLGVETADCEENIESQDRPKLEFFVVERYSWSQLKLIADTRKYHGYMM	60	
Db	1	MAAAMETEQLGVETADCEENIESQDRPKLEFFVVERYSWSQLKLIADTRKYHGYMM	60	
Qy	61	AKAHDPMFVKRNDPDPGPHSDRIYLLAMSGENRENTLFYSEIPKTIINRAAAMLMSKPLL	120	
Db	61	AKAHDPMFVKRNDPDPGPHSDRIYLLAMSGENRENTLFYSEIPKTIINRAAAMLMSKPLL	120	
Qy	121	DLFQATLDYGMYSREBEELLRRKRGITGVGTASVDYHOGSGTFLQAGSGIYHVKGDPQG	180	
Db	121	DLFQATLDYGMYSREBEELLRRKRGITGVGTASVDYHOGSGTFLQAGSGIYHVKGDPQG	180	
Qy	181	FTQQLRNLVETSCPNIRMDPKLCPADPDWIAFIHSDNIWISNIVTRERRLTVVHNEL	240	
Db	181	FTQQLRNLVETSCPNIRMDPKLCPADPDWIAFIHSDNIWISNIVTRERRLTVVHNEL	240	
Qy	241	ANNEEDARSAGVATFVLQEEFDYSGYWMCPKAETTPSGGKILRLIYEENDESEVEI IHV	300	
Db	241	ANNEEDARSAGVATFVLQEEFDYSGYWMCPKAETTPSGGKILRLIYEENDESEVEI IHV	300	
Qy	301	TSPMLETRRADSFYRPTGTANPKVTFKQSEIMIDAEGRIIDVIDKELIQPEILLFEGVE	360	
Db	301	TSPMLETRRADSFYRPTGTANPKVTFKQSEIMIDAEGRIIDVIDKELIQPEILLFEGVE	360	
Qy	361	YIARAGWTPPEKGYAWSILLDRSQTRLQIVLISPELFPVEDDDVMERQRLIESVPDSVTPL	420	
Db	361	YIARAGWTPPEKGYAWSILLDRSQTRLQIVLISPELFPVEDDDVMERQRLIESVPDSVTPL	420	
Qy	421	IYEETTDIWINIHDIHVFPQSHHEEIEFIFASECKTGFRHLYKITSILKESKYGRSSG	480	
Db	421	IYEETTDIWINIHDIHVFPQSHHEEIEFIFASECKTGFRHLYKITSILKESKYGRSSG	480	
Qy	481	GLPAPSDPKCPKEIAITSGEWEVLGRHGSNIQVDEVRRLVYFEGTKQSPLEHLLVYVS	540	
Db	481	GLPAPSDPKCPKEIAITSGEWEVLGRHGSNIQVDEVRRLVYFEGTKQSPLEHLLVYVS	540	
Qy	541	YVNPGEVTRLDRGYSHSCCISQHCDFPI SKYSNQKNPHCVSLYKLSPPEDDPTCKTKGF	600	
Db	541	YVNPGEVTRLDRGYSHSCCISQHCDFPI SKYSNQKNPHCVSLYKLSPPEDDPTCKTKGF	600	
Qy	601	WATILDSAGPLPDYTPPEIFSFESTTGFTLYGMLYKPHDLQPGKYPVTVLFYGGPQVQL	660	
Db	601	WATILDSAGPLPDYTPPEIFSFESTTGFTLYGMLYKPHDLQPGKYPVTVLFYGGPQVQL	660	
Qy	661	VNNRFGVKYFRNLNTLASGLVYVVVVVDNRGSGCHRGKLFEGAPKYKMGQIEIDDOVEGLQY	720	
Db	661	VNNRFGVKYFRNLNTLASGLVYVVVVVDNRGSGCHRGKLFEGAPKYKMGQIEIDDOVEGLQY	720	
Qy	721	LASRYDFIDLDRVGIHGSYGGYLSLMALMQRSDIFRVAIAGAPVTLWIFDGTGYTERYM	780	
Db	721	LASRYDFIDLDRVGIHGSYGGYLSLMALMQRSDIFRVAIAGAPVTLWIFDGTGYTERYM	780	

QY 781 GHPDQEQGYLGSAVMAQAEKFPSEPNRLLLHGLDENVHFAHTSILLSFLVRACKPYD 840
 DB 781 GHPDQEQGYLGSAVMAQAEKFPSEPNRLLLHGLDENVHFAHTSILLSFLVRACKPYD 840
 QY 841 LQIYPERHSIRVPESGEHYELHLLHYLQENLGSRIAALKVI 882
 DB 841 LQIYPERHSIRVPESGEHYELHLLHYLQENLGSRIAALKVI 882

RESULT 5

AAU74749
 ID AAU74749 standard; protein; 882 AA.

XX AAU74749;

XX 09-APR-2002 (first entry)

XX Human protease PR7S-9 protein sequence.

XX Human; protease; PR7S; gastrointestinal; Crohn's disease; cancer;
 KW cardiovascular; atherosclerosis; autoimmune disorder; dermatitis;
 KW inflammatory disorder; acquired immunodeficiency syndrome; AIDS;
 KW cell proliferative disorder; developmental disorder; epilepsy;
 KW Duchenne muscular dystrophy; epithelial disorder; neurological disorder;
 KW reproductive disorder; endometriosis.

XX Homo sapiens.

XX W0200198468-A2.

XX 27-DEC-2001.

XX 13-JUN-2001; 2001WO-US019178.

XX 16-JUN-2000; 2000US-0212336P.

XX 22-JUN-2000; 2000US-0213955P.

XX 29-JUN-2000; 2000US-0215396P.

XX 07-JUL-2000; 2000US-0216821P.

XX 14-JUL-2000; 2000US-0218946P.

XX (INCY-) INCYTE GENOMICS INC.

XX Yue H, Elliott VS, Gandhi AR, Lal P, Au-Young J, Tribouley CM;
 PI Deleagane AM, Baughn MR, Nguyen DB, Lee EA, Hafalia A, Khan FA;
 PI Walia NK, Yao MG, Lu DAM, Patterson C, Tang YT, Walsh RT;
 PI Azimzai Y, Lu Y, Ramkumar J, Xu Y, Reddy R, Das D, Kearney L;
 PI Kallick DA;

XX WPI: 2002-090437/12.

XX N-PSDB; ABK12892.

XX Twenty one human proteases (referred to as PR7S-1 to PR7S-21), useful in
 PT the diagnosis, treatment and prevention of gastrointestinal (e.g.
 PT gastritis), cardiovascular (e.g. atherosclerosis) and cell proliferative
 PT (e.g. cancer) disorders.

XX Claim 1; Page 140-142; 177pp; English.

XX The present invention relates to twenty one new human proteases, referred
 CC to as PR7S-1 to PR7S-21. The PR7S polynucleotides and polypeptides of the
 CC invention are useful in the diagnosis, treatment and prevention of
 CC gastrointestinal e.g. gastritis, esophageal carcinoma and Crohn's
 CC disease, cardiovascular e.g. atherosclerosis, hypertension and myocardial
 CC infarction, autoimmune/inflammatory e.g. acquired immunodeficiency
 CC syndrome (AIDS), allergies and osteoarthritis, cell proliferative e.g.
 CC cancer, developmental e.g. Duchenne and Becker muscular dystrophy,
 CC epithelial e.g. dermatitis, neurological e.g. epilepsy and Alzheimer's
 CC disease and reproductive e.g. infertility and endometriosis disorders.
 CC Numerous other examples of each disorder are given in the specification.
 CC The present protein sequence represents the human protease PR7S-9 protein
 CC of the invention

XX

SQ Sequence 882 AA;

Query Match 100.0%; Score 4700; DB 5; Length 882;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 882; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAAMETEOLGVETPETADCEENIESQDRPKLEPPYVERYSWSOLKKLLADTRKHGYMM 60

DB 1 MAAAMETEOLGVETPETADCEENIESQDRPKLEPPYVERYSWSOLKKLLADTRKHGYMM 60

QY 61 AKAPHDFMFVKRNDPDGPHSDRIYYLAMSNGENRENTLFYSEIPKTIINRAAVALMSWPKLL 120

DB 61 AKAPHDFMFVKRNDPDGPHSDRIYYLAMSNGENRENTLFYSEIPKTIINRAAVALMSWPKLL 120

QY 121 DLFOATLDYGMYSREBELLRERKRGITGVCIASDYDHQSGTFLFOAGSGYHVHVGQPG 180

DB 121 DLFOATLDYGMYSREBELLRERKRGITGVCIASDYDHQSGTFLFOAGSGYHVHVGQPG 180

QY 181 FTQOPLRPNLVETSCPNIRMDPKLCPADPDWIAFIHNSNDIWIISNIVTREERLLTYVHNL 240

DB 181 FTQOPLRPNLVETSCPNIRMDPKLCPADPDWIAFIHNSNDIWIISNIVTREERLLTYVHNL 240

QY 241 ANNEEDARSAGVATFVLQEEFORYGYWPCAKETTPSGGKILRLIYEENDESEVEIHV 300

DB 241 ANNEEDARSAGVATFVLQEEFORYGYWPCAKETTPSGGKILRLIYEENDESEVEIHV 300

QY 301 TSPMLETRRADSPRYPKTGTANPKVTFKMSIIMDAEGRIIDVIDKELIQFELFEGVE 360

DB 301 TSPMLETRRADSPRYPKTGTANPKVTFKMSIIMDAEGRIIDVIDKELIQFELFEGVE 360

QY 361 YIARAGWTPEGKYAWSILLDRSQTRLOIVLISPFLFIVEDDWMERQELLESVPDSVTPL 420

DB 361 YIARAGWTPEGKYAWSILLDRSQTRLOIVLISPFLFIVEDDWMERQELLESVPDSVTPL 420

QY 421 IIVEETTDWINIHDFHVPQSHHEEIEFIPASCKTGRHLKYITSILKESYKSSG 480

DB 421 IIVEETTDWINIHDFHVPQSHHEEIEFIPASCKTGRHLKYITSILKESYKSSG 480

QY 481 GLPAPSDFKCPKEEIAITSGEWEVLGRHGSNIQVDEVRRLVYFEGTKDSPLEHLLYVVS 540

DB 481 GLPAPSDFKCPKEEIAITSGEWEVLGRHGSNIQVDEVRRLVYFEGTKDSPLEHLLYVVS 540

QY 541 YNPGEVTRLTDRGYSHSCCISQHCDFPISKYSNQKNPHCVSLKLSPPEDDPTCKTKEF 600

DB 541 YNPGEVTRLTDRGYSHSCCISQHCDFPISKYSNQKNPHCVSLKLSPPEDDPTCKTKEF 600

QY 601 WATILDSAGPLPDYTPPEIFSFESTTGTLYGMLYKPHDLQPKKYPTVLFIYGGPOVOL 660

DB 601 WATILDSAGPLPDYTPPEIFSFESTTGTLYGMLYKPHDLQPKKYPTVLFIYGGPOVOL 660

QY 661 VNNRFKGVKVFRLNTLASLGYVWVVDNRGSHRGKFEKAFKYMKGQIEIDDOVEGLQY 720

DB 661 VNNRFKGVKVFRLNTLASLGYVWVVDNRGSHRGKFEKAFKYMKGQIEIDDOVEGLQY 720

QY 721 LASRYDFIDLDRVGIHGSYGGYLSMALMQRSDIFRVAIAGAPVTLWIFVDTGYTRYM 780

DB 721 LASRYDFIDLDRVGIHGSYGGYLSMALMQRSDIFRVAIAGAPVTLWIFVDTGYTRYM 780

QY 781 GHPDQEQGYLGSAVMAQAEKFPSEPNRLLLHGLDENVHFAHTSILLSFLVRACKPYD 840

DB 781 GHPDQEQGYLGSAVMAQAEKFPSEPNRLLLHGLDENVHFAHTSILLSFLVRACKPYD 840

QY 841 LQIYPERHSIRVPESGEHYELHLLHYLQENLGSRIAALKVI 882

DB 841 LQIYPERHSIRVPESGEHYELHLLHYLQENLGSRIAALKVI 882

RESULT 6

AD117086

ID AD117086 standard; protein; 882 AA.

XX AC

XX AD117086;

XX XX

DT	15-APR-2004	(first entry)	PA	(CURA-) CURAGEN CORP.
XX			XX	Tchernev VT, Spytek KA, Zerhusen BD, Patturajan M, Shimkets RA;
DE		Human NOVX protein homologue seqID 622.	PI	Li L, Gangolli EA, Padigaru M, Anderson DM, Rastelli L, Miller CE;
XX		human; NOVX; cardiomyopathy; atherosclerosis; cancer; diabetes;	PI	Gerlach VL, Taupier RU, Gusev VY, Colman SD, Wolenc AR, Pena CEA;
KW		inflammation; autoimmune disorder; allergy; blood disorder;	PI	Furtak K, Grosse WM, Alsobrook JP, Lepley DM, Rieger DK, Burgess CE;
KW		acquired immunodeficiency syndrome; AIDS; obesity; asthma;	XX	WPI; 2002-706998/76.
KW		immunoglobulin (Ig)A nephropathy; cirrhosis; arthritis;	XX	New NOVX polypeptides and nucleic acids, useful for preventing or
KW		Alzheimer's disease; infection; str.	PT	treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,
OS		Homo sapiens.	PT	atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or
XX		WO200268649-A2.	PT	pharmacogenomics.
PN		06-SEP-2002.	XX	Disclosure; SEQ ID NO 622; 1498pp; English.
XX			XX	This invention relates to a novel nucleic acids, and encoded polypeptides
XX	31-JAN-2002;	2002WO-US002785.	CC	thereof, which have properties related to the stimulation of biochemical
XX			CC	or physiological responses in a cell, tissue, organ or organism.
PR	31-JAN-2001;	2001US-0265395P.	CC	Specifically, it refers to the use of biologically active fragments for
PR	31-JAN-2001;	2001US-0265514P.	CC	diagnostic and prognostic assays and furthermore in the treatment of
PR	31-JAN-2001;	2001US-0265517P.	CC	diverse pathological conditions. The present invention describes novel
PR	05-FEB-2001;	2001US-0266406P.	CC	human and murine NOVX proteins, as well as methods to modulate their
PR	02-FEB-2001;	2001US-0266767P.	CC	expression using antisense oligos, ribozymes and peptide nucleic acids.
PR	07-FEB-2001;	2001US-0266975P.	CC	The NOVX polypeptides, polynucleotides and antibodies are useful in
PR	07-FEB-2001;	2001US-0267057P.	CC	treating or preventing NOVX-associated disorders, e.g. cardiomyopathy,
PR	08-FEB-2001;	2001US-0267459P.	CC	atherosclerosis, cancer and diabetes. Furthermore, they may be used in
PR	09-FEB-2001;	2001US-0267823P.	CC	treating or preventing diseases such as inflammation, autoimmune
PR	15-FEB-2001;	2001US-0268974P.	CC	disorders, allergies, asthma, immunoglobulin (Ig)A nephropathy, cirrhosis,
PR	26-FEB-2001;	2001US-0271664P.	CC	(AIDS), obesity, stroke, infections, stroke, muscular dystrophy,
PR	27-FEB-2001;	2001US-0271839P.	CC	arthritis, Alzheimer's disease, immunosuppressive, antiallergic,
PR	27-FEB-2001;	2001US-0271855P.	CC	cytostatic, cardiant, antidiabetic, antiarteriosclerotic, anorectic,
PR	02-MAR-2001;	2001US-0272046P.	CC	haemostatic, anti-HIV, antiinflammatory, antitumor, antiparasitic,
PR	02-MAR-2001;	2001US-0275925P.	CC	antiparasitic, neurotropic, antihypertensive, antidiabetic,
PR	14-MAR-2001;	2001US-0275947P.	CC	neuroprotective, neurotropic, antihypertensive, antidiabetic,
PR	14-MAR-2001;	2001US-0275950P.	CC	relaxant and anticonvulsant. In addition, they are useful in screening
PR	14-MAR-2001;	2001US-0275980P.	CC	assays to identify small molecules that modulate or inhibit, for example,
PR	14-MAR-2001;	2001US-0275982P.	CC	neurogenesis, wound healing and angiogenesis. The nucleic acids are also
PR	15-MAR-2001;	2001US-0276448P.	CC	used as in chromosome mapping, tissue typing, preventive medicine and
PR	15-MAR-2001;	2001US-0276450P.	CC	pharmacogenomics. This polypeptide is a homologue of a human NOVX protein
PR	16-MAR-2001;	2001US-0276597P.	CC	of the invention.
PR	16-MAR-2001;	2001US-0276598P.	XX	Sequence 882 AA;
PR	20-MAR-2001;	2001US-0278652P.	XX	Query Match 100.0%; Score 4700; DB 5; Length 882;
PR	26-MAR-2001;	2001US-0278775P.	XX	Best Local Similarity 100.0%; Pred. No. 0;
PR	26-MAR-2001;	2001US-0278778P.	XX	Matches 882; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
PR	29-MAR-2001;	2001US-0279682P.		
PR	29-MAR-2001;	2001US-0279684P.		
PR	30-MAR-2001;	2001US-0280147P.		
PR	11-APR-2001;	2001US-0282992P.		
PR	11-APR-2001;	2001US-0283083P.		
PR	20-APR-2001;	2001US-0285133P.		
PR	23-APR-2001;	2001US-0285749P.		
PR	03-MAY-2001;	2001US-0288327P.		
PR	03-MAY-2001;	2001US-0288504P.		
PR	29-MAY-2001;	2001US-0294047P.		
PR	30-MAY-2001;	2001US-0294473P.		
PR	08-JUN-2001;	2001US-0296964P.		
PR	18-JUN-2001;	2001US-0298959P.		
PR	19-JUN-2001;	2001US-0299324P.		
PR	13-AUG-2001;	2001US-0312020P.		
PR	16-AUG-2001;	2001US-0312889P.		
PR	16-AUG-2001;	2001US-0312908P.		
PR	21-AUG-2001;	2001US-0313390P.		
PR	28-AUG-2001;	2001US-0315470P.		
PR	31-AUG-2001;	2001US-0316447P.		
PR	07-SEP-2001;	2001US-0318115P.		
PR	07-SEP-2001;	2001US-0318118P.		
PR	12-SEP-2001;	2001US-0318740P.		
PR	19-SEP-2001;	2001US-0323379P.		
PR	18-OCT-2001;	2001US-0330245P.		
PR	18-OCT-2001;	2001US-0330308P.		
PR	14-NOV-2001;	2001US-0332701P.		
XX				
Qy	1	MAAAMETEQLGVEIFETADCEENIESQDRPKLEPFYVERYSWSQKLLADTRKYHYMM	60	
Db	1	MAAAMETEQLGVEIFETADCEENIESQDRPKLEPFYVERYSWSQKLLADTRKYHYMM	60	
Qy	61	AKAPHPMFVKNDPDGPHSDRIYVLAMSGENRENTLFYSEIPKTIINRAAVLMSWKPLL	120	
Db	61	AKAPHPMFVKNDPDGPHSDRIYVLAMSGENRENTLFYSEIPKTIINRAAVLMSWKPLL	120	
Qy	121	DLFOATLDYGMYSREELRLERKRTGTGVIASDYHOGSGTFLFOAGSGYHVHKGDPQG	180	
Db	121	DLFOATLDYGMYSREELRLERKRTGTGVIASDYHOGSGTFLFOAGSGYHVHKGDPQG	180	
Qy	181	FTQOPLRPNLVETSCNIRMDPKLCPADPDWIAFHNSNDIWSNIVTFRERLLTYVHNL	240	
Db	181	FTQOPLRPNLVETSCNIRMDPKLCPADPDWIAFHNSNDIWSNIVTFRERLLTYVHNL	240	
Qy	241	ANMEEDARSAGVATFVLQEFDRYSGYWWCPKAEPTTPSGGKLILILYEENDESEVLIHV	300	
Db	241	ANMEEDARSAGVATFVLQEFDRYSGYWWCPKAEPTTPSGGKLILILYEENDESEVLIHV	300	
Qy	301	TSPMLETRRADSPRYPKTGTANPKVTFKMSIIMDAEGRIIDVIDKELIQFELLFGVE	360	
Db	301	TSPMLETRRADSPRYPKTGTANPKVTFKMSIIMDAEGRIIDVIDKELIQFELLFGVE	360	

QY 361 YIARAGWTPEKGYAWSILLDRSQRLOIQLVLSPELFIPIVEDDWMERQRLIESVDPSTVPL 420
DB |||||
QY 361 YIARAGWTPEKGYAWSILLDRSQRLOIQLVLSPELFIPIVEDDWMERQRLIESVDPSTVPL 420
DB |||||
QY 421 ILYEETTDIWINIHDPVPOSHHEIEFIFASECKTGFRHLKYITSLILKESKYKSSG 480
DB |||||
QY 421 ILYEETTDIWINIHDPVPOSHHEIEFIFASECKTGFRHLKYITSLILKESKYKSSG 480
DB |||||
QY 481 GLPAPDFKCPKEEIAITSGEWEVLGRHGSNIQVDEVRRLVYFEGTKDSPLEHLLVVS 540
DB |||||
QY 481 GLPAPDFKCPKEEIAITSGEWEVLGRHGSNIQVDEVRRLVYFEGTKDSPLEHLLVVS 540
DB |||||
QY 541 YVNPGEVTRLTDRGYSHSCCISQCHDFFIKYSQKNPHCVSLYKLSPPEDDPTCKTKEF 600
DB |||||
QY 541 YVNPGEVTRLTDRGYSHSCCISQCHDFFIKYSQKNPHCVSLYKLSPPEDDPTCKTKEF 600
DB |||||
QY 601 WATILDSAGPLDPTYPEIFSFSTTGFTLYGMLYKPHDLQPGKKYPTVLFYIGGPOVQL 660
DB |||||
QY 601 WATILDSAGPLDPTYPEIFSFSTTGFTLYGMLYKPHDLQPGKKYPTVLFYIGGPOVQL 660
DB |||||
QY 661 VNNRFGVKYFRLNTLASLYGVVVVINDNRGSHRGKLPKFGAPKYKMGQIIDDQVEGLQY 720
DB |||||
QY 661 VNNRFGVKYFRLNTLASLYGVVVVINDNRGSHRGKLPKFGAPKYKMGQIIDDQVEGLQY 720
DB |||||
QY 721 LASRYDFIDLRVGIHWSYGGYLSLMLMQRSDIFRVAIAGAPVTLWIFDGTGYTRYM 780
DB |||||
QY 721 LASRYDFIDLRVGIHWSYGGYLSLMLMQRSDIFRVAIAGAPVTLWIFDGTGYTRYM 780
DB |||||
QY 781 GHPPONQGYVLSVAMQAEKFPSEPNRLLHGFLENVHFAHTSILLSLFLVRAGKPYD 840
DB |||||
QY 781 GHPPONQGYVLSVAMQAEKFPSEPNRLLHGFLENVHFAHTSILLSLFLVRAGKPYD 840
DB |||||
QY 841 LQIYQERHSIRVPESGEHYELHLLHYLOENLGSRIAALKVI 882
DB |||||
QY 841 LQIYQERHSIRVPESGEHYELHLLHYLOENLGSRIAALKVI 882
DB |||||

RESULT 7

ABU07720
ID ABU07720 standard; protein; 882 AA.
XX AC ABU07720;
XX DT 19-MAY-2003 (first entry)
XX DE Human serine protease HIPHUM46.
XX KW Human; enzyme; HIPHUM46; serine protease; gene therapy; osteoarthritis;
KW serine protease activity modulation; dipeptidyl peptidase activity;
KW musculoskeletal disease; Hepatitis B virus infection; myotonic dystrophy;
KW Alzheimer's disease; parasupranuclear palsy; Huntington's disease;
KW amyotrophic lateral sclerosis; malabsorption syndrome; lung disease;
KW irritable bowel syndrome; type I diabetes; faecal incontinence; anaemia;
KW haemorrhoid; proctitis; rectal polyp; small bowel tumour; dyslexia;
KW colorectal tumour; ceroid lipofuscinosis; allergic encephalomyelitis;
KW multiple sclerosis.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
FT Region 259..260
FT /note= "Paired glutamates of the beta propeller domain"
FT Active-site 739
FT /label= Catalytic_serine_residue
FT Active-site 817
FT /label= Catalytic_aspartate_residue
FT Active-site 849
FT /label= Catalytic_histidine_residue
XX GB2374869-A.
XX PN
XX PD 30-OCT-2002.
XX

PF 22-JAN-2002; 2002GB-00001404.
XX 23-JAN-2001; 2001GB-00001760.
XX (GLAX) GLAXO GROUP LTD.
XX Edbrooke MR, Lewis AP;
XX WPI; 2003-150703/15.
XX N-PSDB; ABX12255.
PT Identifying modulators of serine protease activity useful for treating
PT musculoskeletal diseases, by contacting cell expressing a novel serine
PT protease polypeptide with a compound and monitoring serine protease
PT activity.
XX Claim 10; Page 26-29; 38pp; English.
XX The invention relates to a method of identifying a substance that
CC modulates serine protease activity, comprising contacting a cell such as
CC a neuronal cell, lung cell, intestinal cell or a cell infected with a
CC virus, expressing a serine protease polypeptide (HIPHUM 46), or its
CC variant having dipeptidyl peptidase activity, or a serine protease
CC isolated from the cell with a test substance and monitoring for serine
CC protease activity. The method is useful for identifying a substance that
CC modulates serine protease activity. A modulator of the serine protease is
CC useful in the manufacture of a medicament for treatment or prophylaxis of
CC a musculoskeletal disease e.g. osteoarthritis, Hepatitis B virus
CC infection, Alzheimer's disease, parasupranuclear palsy, myotonic
CC dystrophy, Huntington's disease or amyotrophic lateral sclerosis.
CC Additional disease that may be treated using modulators of the serine
CC protease include malabsorption syndromes, irritable bowel syndrome, lung
CC disease, type I diabetes, faecal incontinence, haemorrhoids, proctitis,
CC rectal polyps, small bowel tumours, colorectal tumours, anaemia,
CC dyslexia, ceroid lipofuscinosis, allergic encephalomyelitis, and multiple
CC sclerosis. The present sequence represents the amino acid sequence of the
CC human serine protease HIPHUM46
XX
SQ Sequence 882 AA;
Query Match 100.0%; Score 4700; DB 6; Length 882;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 882; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAAMETEQLGVIEFETADCEENIESQDRPKLEPFVVERYSWSQLKLADTRKYHYGM 60
DB |||||
QY 1 MAAMETEQLGVIEFETADCEENIESQDRPKLEPFVVERYSWSQLKLADTRKYHYGM 60
DB |||||
QY 61 AKAPDPMFKVRNDPDGPHSDRIYYLAMSGENRENTLFYSEIPKTNRAAAMLWSKPL 120
DB |||||
QY 61 AKAPDPMFKVRNDPDGPHSDRIYYLAMSGENRENTLFYSEIPKTNRAAAMLWSKPL 120
DB |||||
QY 121 DLFOATLDYGMYSREBELLRERKRICTVGIASYDYHQSGTFLFOAGSGIYHVDGPGQ 180
DB |||||
QY 121 DLFOATLDYGMYSREBELLRERKRICTVGIASYDYHQSGTFLFOAGSGIYHVDGPGQ 180
DB |||||
QY 181 FTQQLRPNLVETSCPNIRMDPKLCPADPDWIAFIHSDNIWISNIVTREERRLTYVHNEL 240
DB |||||
QY 181 FTQQLRPNLVETSCPNIRMDPKLCPADPDWIAFIHSDNIWISNIVTREERRLTYVHNEL 240
DB |||||
QY 241 ANMEEDARSAGVATFVLQEEFDYSGYWWCPKAETTPSGGKILRLIYENDESEVEIHHV 300
DB |||||
QY 241 ANMEEDARSAGVATFVLQEEFDYSGYWWCPKAETTPSGGKILRLIYENDESEVEIHHV 300
DB |||||
QY 301 TSPMLETRRADSFYRKPTGTANPKVTFKMSIEMI DAEGRIIDVIDKELIQPEILFEVGE 360
DB |||||
QY 301 TSPMLETRRADSFYRKPTGTANPKVTFKMSIEMI DAEGRIIDVIDKELIQPEILFEVGE 360
DB |||||
QY 361 YIARAGWTPEKGYAWSILLDRSQRLOIQLVLSPELFIPIVEDDWMERQRLIESVDPSTVPL 420
DB |||||
QY 361 YIARAGWTPEKGYAWSILLDRSQRLOIQLVLSPELFIPIVEDDWMERQRLIESVDPSTVPL 420
DB |||||
QY 421 ILYEETTDIWINIHDPVPOSHHEIEFIFASECKTGFRHLKYITSLILKESKYKSSG 480
DB |||||

CC neuroprotective, neurotropic, antibacterial, virucide, antiparasitic,
CC relaxant and anticonvulsant. In addition, they are useful in screening
CC assays to identify small molecules that modulate or inhibit, for example,
CC neurogenesis, wound healing and angiogenesis. The nucleic acids are also
CC used as in chromosome mapping, tissue typing, preventive medicine and
CC pharmacogenomics. This polypeptide is a homologue of a human NOVX protein
CC of the invention.

Sequence 883 AA;

Query Match	96.4%;	Score 4528.5;	DB 5;	Length 883;
Best Local Similarity	95.7%;	Pred. No. 0;		
Matches 845;	Conservative 18;	Mismatches 19;	Indels 1;	Gaps 1;
Qy	1	MAAAMETGOLGVEIIFETADCBEE-NIESQDRPKLEPFVVERYSWSQLKLLADTRKYHGVM	59	
Db	1	MAAAMETGOLGVEIIFETADCBEE-NIESQDRPKLEPFVVERYSWSQLKLLADTRKYHGVM	60	
Qy	60	MAKAPHDFMFKVRNDPDGPHSDRIYYLAMSGENRENTLFYSEIPKTNRAAVALMLSWKPL	119	
Db	61	MAKAPHDFMFKVRKTPDRPHSDRVYVYLAMSGENRENTLFYSEIPKTNRAAVALMLSWKPL	120	
Qy	120	LDLFOATLDYGYMSREBELLRKRIIGTVGTIASVDYHQSGETFLFOAGSGIYHVKGQGP	179	
Db	121	LDLFOATLDYGYMSREBELLRKRIIGTVGTIAAYHGPSGETFLFOAGSGIYHIXDGGPH	180	
Qy	180	GFTQQPLRNLVETSCPNIRMDPKLCPADPDWIAFIHSNDIWSINVTREERLTYVNE	233	
Db	181	GFTQQPLRNLVETSCPNIRMDPKLCPADPDWIAFIHSNDIWSINVTREERLTYVNE	240	
Qy	240	LANNEEDARSAGVAIFVLQEEFDRYSGYWCCKAETTPSGGKILRLIYBENDESEVEIHH	299	
Db	241	LANNEEDPRSAGVAIFVLQEEFDRYSGYWCCKAETTPSGGKILRLIYBENDESEVEIHH	300	
Qy	300	VTSPLMETRRASFRYPKTGTANPKVTPKASEIMIDAEGRIOVDIKELIQPEILFGV	359	
Db	301	VTSPLMETRRASFRYPKTGTANPKVTPKMSIEIIVDAAGGIOVDIKELIQPEILFGV	360	
Qy	360	EYARAGWTPGEGKAWLSILLDRSOTRLOI VLI SPELFI PVEDDWERQRLIESVPDSTVP	419	
Db	361	EYARAGWTPGEGKAWLSILLDRSQTHLOI VLI SPELFI PVEDDAMQRLIESVPDSTVP	420	
Qy	420	LIYYEETDIWINIHDI FHVFPQSHEEIEIFASECKTGFRLHYKITSLKESKYGRSS	479	
Db	421	LIYYEETDIWINIHDI FHVFPQTHEDEIEIFASECKTGFRLHYKITSLKESKYGRSS	480	
Qy	480	GGLPAPSDFKPCIKESIAITSGEWEVLGRHGSNTQDVDERRLVYFEGTKDSPLEHLLYV	539	
Db	481	GGLPAPSDFKPCIKEITITSGEWEVLGRHGSNTWDEARKLYFEGTKDSPLEHLLYV	540	
Qy	540	SYVNPGEVRLTRDGRYSHSCCLSOHCDFFI SKYSNQKNPHCVSIYKLSPEDDPTCKTKE	599	
Db	541	SYANPGEVRLTRDGRYSHSCCLSRHCDFFI SKYSNQKNPHCVSIYKLSPEDDPVHKTKE	600	
Qy	600	FWATILDSAGLPDYTPPEIFSFESTTGFTLYGMLYKPHDLQPKKYPTVLFYIGGPQVQ	659	
Db	601	FWATILDSAGLPDYTPPEIFSFESTTGFTLYGMLYKPHDLQPKKYPTVLFYIGGPQVQ	660	
Qy	660	LVNRRFGVKYPLRNLASLGVVVVVVDNRGSCHRGLKPFEGAFYKMGQIIEDDQVEGLQ	719	
Db	661	LVNRRFGVKYPLRNLASLGVVVVVVDNRGSCHRGLKPFEGAFYKMGQIIEDDQVEGLQ	720	
Qy	720	YLASRVDFDLDVRVGIHGSYGGYLSIALMQRSDIFRVAITAGAPVTLMIFYDTGVTERY	779	
Db	721	YLASQVDFDLDVRVGIHGSYGGYLSIALMQRSDIFRVAITAGAPVTLMIFYDTGVTERY	780	
Qy	780	MGPDPQNEQGYVILGYSVAMQAEKFPSEPNRLILLHGFLDENVHFAHTSILLSFVLRAKGPY	839	
Db	781	MGPDPQNEQGYVILGYSVAMQAEKFPSEPNRLILLHGFLDENVHFAHTSILLSFVLRAKGPY	840	
Qy	840	DLQIYPOERHSIRVPESGHEYLHLLHYLOENLGSRTAAUKVI	882	
Db	841	DLQIYPOERHSIRVPESGHEYLHLLHYLOENLGSRTAAUKVI	883	

DT	XX	DE	XX	Human protein modification and maintenance molecule-10 (PMWM-10).
15-JUL-2003	(first entry)			
Human protein modification and maintenance molecule-10 (PMWM-10).				
Human; protein modification and maintenance molecule; PMWM; cancer; cell proliferation disorder; atherosclerosis; neurological disorder; epilepsy; Huntington's disease; stroke; immune disorder; allergy; inflammatory disorder; AIDS; developmental disorder; hypothyroidism; Cushing's syndrome; gastrointestinal disorder; epithelial disorder; infection; cytotoxic; antiarthrosclerotic; anticonvulsant; nootropic; neuroprotective; cerebroprotective; anti-HIV; antiallergic; vulnary; antiinflammatory; thymomimetic.				
Homo sapiens.				
WO2003031939-A2.				
17-APR-2003.				
11-OCT-2002; 2002WO-US032850.				
12-OCT-2001; 2001US-0329689P.				
25-OCT-2001; 2001US-0335703P.				
09-NOV-2001; 2001US-0348887P.				
28-NOV-2001; 2001US-0334145P.				
06-DEC-2001; 2001US-0337451P.				
14-DEC-2001; 2001US-0340584P.				
(INCY-) INCYTE GENOMICS INC.				
Ramkumar J, Gorvad AE, Baughn MR, Emerling BM, Yang J, Lee SY, Tran UK, Becha SD, Duggan BM, Lee EA, Griffin JA, Li, Sprague W, Hafalia AJA, Chawla NK, Lehr-Mason PM, Kable AE, Yue H, Marquis JP, Yao MG, Richardson TW, Tang TV, Jin P, Chien D, Bhadia U, Burrill JD, Lee S, Blake JJ, Ho A, Zheng W, NPI; 2003-430274/40.				
DR	XX	DR	XX	PMWSDB; ACA9425.
New human protein modification and maintenance molecules (PMWM), useful for diagnosing, treating and preventing diseases or conditions associated with the aberrant PMWM expression e.g. cancer, atherosclerosis, or infections.				
Claim 1; Page 249-251; 31pp; English.				
The present invention relates to the isolation of human protein modification and maintenance molecules (PMWM), and the polynucleotide sequences encoding them. A total of 40 PMWM polypeptides (designated PMWM-1 to PMWM-40) are disclosed. The sequences of the invention are useful for diagnosing a condition or disease associated with the expression of PMWM in a subject, preparing a polyclonal or monoclonal antibody, and generating an expression profile of a sample containing the polynucleotides. The diseases or conditions associated with decreased expression or overexpression of PMWM are cell proliferation disorders (e.g. cancer, atherosclerosis), neurological disorders (e.g. epilepsy, Huntington's disease, stroke), immune/inflammatory disorders, Cushing's syndrome, developmental disorders (e.g. hypothyroidism, Cushing's syndrome), gastrointestinal or epithelial disorders, and infections. The PMWM polypeptides or their fragments are useful in screening compounds for effectiveness as agonists or antagonists of the polypeptides, or in altering the expression of the target polynucleotide and compounds that specifically bind to, or modulate the activity of the polypeptide.				
ABU92021-ABU92060 represent the human PMWM polypeptides of the invention				
Sequence 824 AA;				
Query Match	92.6%;	Score	4353;	DB 6; Length 824;
Best Local Similarity	93.4%;	Pred	No. 0;	
Matches	824;	Conservative	0;	Mismatches
0;				Indels
58;				Gaps
1				
MAAAMETQLGVEIFETADCEENIESQDRPKLEPFYVERYSQKLADTRKYHYGM				
60				
1				
MAAAMETQLGVEIFETADCEENIESQDRPKLEPFYVERYSQKLADTRKYHYGM				
60				
61				
AKAPHDFMFKVRNDPDPGPHSDRIYYLAMSGENRENTLFYSEIPKTIINRAA				
120				
61				
AKAPHDFMFKVRNDPDPGPHSDRIYYLAMSGENRENTLFYSEIPKTIINRAA				
120				
121				
DLFOATLDYGMYSREBELLRKRIGTVGIASDYHQSGTFLFQAGSGIYHV				
180				
121				
DLFOATLDYGMYSREBELLRKRIGTVGIASDYHQSGTFLFQAGSGIYHV				
180				
181				
FTQOPLRNLVETSCPNIRMDPKLCPADPDWIAFIHNSNDIWTISNIVT				
240				
181				
FTQOPLRNLVETSCPNIRMDPKLCPADPDWIAFIHNSNDIWTISNIVT				
240				
241				
ANMEDARSAGVAFVLOEPEDRYSGVWCPKAEFTPSGKILRLIYENDESE				
300				
241				
ANMEDARSAGVAFVLOEPEDRYSGVWCPKAEFTPSGKILRLIYENDESE				
300				
301				
TSPMLETRADSFRYPKGTGTANPKVTFQSEIMIDAEGRIDVIDKELI				
360				
301				
TSPMLETRADSFRYPKGTGTANPKVTFQSEIMIDAEGRIDVIDKELI				
360				
361				
YIARAGWTPEGKYAWSILLDRSQTRLOIVLISPELFIPIVEDDVMERQ				
420				
361				
YIARAGWTPEGKYAWSILLDRSQTRLOIVLISPELFIPIVEDDVMERQ				
420				
421				
IIYEETDIWINIHDIHFVPPQSHEEIEFIFASECKTGRHLKYITSLKES				
480				
421				
IIYEETDIWINIHDIHFVPPQSHEEIEFIFASECKTGRHLKYITSLKES				
480				
481				
GLPAPSDPKPIKEBIAITSGEWEVLGRHGSNIQVDEVRRLVYFEGTK				
540				
481				

Db 1 MAAAMETEQLGVEIFETADCEENIESQDRPKLEPPYVERYSWSQKLLADTRKYGYMM 60
 Qy 61 AKAPHPDMFKVRNDPGPHSDRIYYLAMSNGENRENTLFYSEIPKTIINRAAVLMLSWKPL 120
 Db 61 AKAPHPDMFKVRNDPGPHSDRIYYLAMSNGENRENTLFYSEIPKTIINRAAVLMLSWKPL 120
 Qy 121 DLFOATLDYGMYSREBELLRERKRIGTVGIASDYHQSGTFLFOAGSGIYHVKGDPQG 180
 Db 121 DLFO----- 124
 Qy 181 FTQOPLRPNLVETSCNIRMDPKLCPADPDWIAFTHSNDIWIISNIVTREERLLTYVHNEL 240
 Db 125 --OOLRPNLVETSCNIRMDPKLCPADPDWIAFTHSNDIWIISNIVTREERLLTYVHNEL 182
 Qy 241 ANMEEDARAGVATFVLQEFDRYSGYWCPCAEITPSSGGKILRLIYEENDESEVEIHHV 300
 Db 183 ANMEEDARAGVATFVLQEFDRYSGYWCPCAEITPSSGGKILRLIYEENDESEVEIHHV 242
 Qy 301 TSPMLETRRADSPRYPKTGTANPKVTFKMSIEMIDAEGRIIDVIDKELIQPFELFEGVE 360
 Db 243 TSPMLETRRADSPRYPKTGTANPKVTFKMSIEMIDAEGRIIDVIDKELIQPFELFEGVE 302
 Qy 361 YIARAGWTPGKYAWSILLDRSQTRLOIQLVISPFLIPVEDDVMERQRLIESVPDSVTP 420
 Db 303 YIARAGWTPGKYAWSILLDRSQTRLOIQLVISPFLIPVEDDVMERQRLIESVPDSVTP 362
 Qy 421 IIVEETTDIWINIHDFHVPQSHHEEIEFIPASECKTGFRHLKYITSLKESKYRSSG 480
 Db 363 IIVEETTDIWINIHDFHVPQSHHEEIEFIPASECKTGFRHLKYITSLKESKYRSSG 422
 Qy 481 GLPAPSDFKPIKEEIAITSGWEVLGRHGSNIQVDEVRRLVYFEGTKDSPLEHLYVVS 540
 Db 423 GLPAPSDFKPIKEEIAITSGWEVLGRHGSNIQVDEVRRLVYFEGTKDSPLEHLYVVS 482
 Qy 541 YNPGEVTRLTDGYSHSCCISQHCDFPISKYNSQKPHCVSLYKLSPEDDPTCKTKEF 600
 Db 483 YNPGEVTRLTDGYSHSCCISQHCDFPISKYNSQKPHCVSLYKLSPEDDPTCKTKEF 542
 Qy 601 WATILDSAGPLDPTPEIFSPFESTTGTLYGMLYKPHDLOPKKYPTVLFIYGGPOVL 660
 Db 543 WATILDSAGPLDPTPEIFSPFESTTGTLYGMLYKPHDLOPKKYPTVLFIYGGPOVL 602
 Qy 661 VNNRFKGVKFRNLTLASGYVWVVDNRGSHRGKLEKPEGAFKMGQIEIDQVEGLQY 720
 Db 603 VNNRFKGVKFRNLTLASGYVWVVDNRGSHRGKLEKPEGAFKMGQIEIDQVEGLQY 662
 Qy 721 LASRYDFIDLDRVGIHWSYGYLSLMLMORSIDIFRVAIAGAPVTLWIFDYGTYERYM 780
 Db 663 LASRYDFIDLDRVGIHWSYGYLSLMLMORSIDIFRVAIAGAPVTLWIFDYGTYERYM 722
 Qy 781 GHPDNEQGYLGSVAMQAEKPSBNRLLLHGFLENVHFAHTSILLSPFLVRAGKPYD 840
 Db 723 GHPDNEQGYLGSVAMQAEKPSBNRLLLHGFLENVHFAHTSILLSPFLVRAGKPYD 782
 Qy 841 LOIYPQERHSIRVPESGEHYELHLHYLOENLGSRIIAALKVI 882
 Db 783 LOIYPQERHSIRVPESGEHYELHLHYLOENLGSRIIAALKVI 824

RESULT 11

ABB97361 ID ABB97361 standard; protein; 782 AA.

XX AC ABB97361;

XX DT 27-JUN-2002 (first entry)

XX DE Novel human protein SEQ ID NO: 629.

XX KW Human; antianaemic; vulnary; antiinflammatory; immunomodulator;
 KW antiinfertility; cerebroprotective; cytostatic; rheumatic; gene therapy;
 KW neuroprotective; antiparkinsonian; protein therapy; BST;
 KW expressed sequence tag.

XX OS Homo sapiens.
 XX PN WO200222660-A2.
 XX PD 21-MAR-2002.
 XX PF 10-SEP-2001; 2001WO-US026015.
 XX PR 11-SEP-2000; 2000US-00659671.
 XX PA (HYSE-) HYSEQ INC.
 XX PI Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
 PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;
 XX WPI: 2002-292408/33.
 XX DR N-PSDB; ABB32547.
 XX PT An isolated polynucleotide for treating diseases associated with its
 XX encoded polypeptide such as cancer and multiple sclerosis.
 XX PS Example 2; SEQ ID NO 629; 509pp; English.
 XX CC The present invention provides the protein and coding sequences of 444
 CC novel human proteins. These were isolated from expressed sequences tags
 CC (ESTs). They can be used to stimulate cell growth, to regulate
 CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
 CC e.g. in burn treatment, to regulate the immune system e.g. to treat
 CC multiple sclerosis, to regulate activin or inhibin e.g. to treat
 CC infertility, to regulate haemostasis or thrombolysis e.g. to treat stroke
 CC and cancer, to screen for drugs, to treat inflammatory conditions e.g.
 CC rheumatoid arthritis, and to treat nervous system disorders e.g.
 CC Parkinson's disease. The present sequence is a protein of the invention
 XX Sequence 782 AA;

Query Match 87.6%; Score 4118; DB 5; Length 782;
 Best Local Similarity 88.7%; Pred. No. 0;
 Matches 782; Conservative 0; Mismatches 0; Indels 100; Gaps 1;

Qy 1 MAAAMETEQLGVEIFETADCEENIESQDRPKLEPPYVERYSWSQKLLADTRKYGYMM 60
 Db 1 MAAAMETEQLGVEIFETADCEENIESQDRPKLEPPYVERYSWSQKLLADTRKYGYMM 60
 Qy 61 AKAPHPDMFKVRNDPGPHSDRIYYLAMSNGENRENTLFYSEIPKTIINRAAVLMLSWKPL 120
 Db 61 AKAPHPDMFKVRNDPGPHSDRIYYLAMSNGENRENTLFYSEIPKTIINRAAVLMLSWKPL 120
 Qy 121 DLFOATLDYGMYSREBELLRERKRIGTVGIASDYHQSGTFLFOAGSGIYHVKGDPQG 180
 Db 121 DLFOATLDYGMYSREBELLRERKRIGTVGIASDYHQSGTFLFOAGSGIYHVKGDPQG 180
 Qy 181 FTQOPLRPNLVETSCNIRMDPKLCPADPDWIAFTHSNDIWIISNIVTREERLLTYVHNEL 240
 Db 181 FTQOPLRPNLVETSCNIRMDPKLCPADPDWIAFTHSNDIWIISNIVTREERLLTYVHNEL 240
 Qy 241 ANMEEDARAGVATFVLQEFDRYSGYWCPCAEITPSSGGKILRLIYEENDESEVEIHHV 300
 Db 241 ANMEEDARAGVATFVLQEFDRYSGYWCPCAEITPSSGGKILRLIYEENDESEVEIHHV 300
 Qy 301 TSPMLETRRADSPRYPKTGTANPKVTFKMSIEMIDAEGRIIDVIDKELIQPFELFEGVE 360
 Db 301 TSPMLETRRADSPRYPKTGTANPKVTFKMSIEMIDAEGRIIDVIDKELIQPFELFEGVE 360
 Qy 361 YIARAGWTPGKYAWSILLDRSQTRLOIQLVISPFLIPVEDDVMERQRLIESVPDSVTP 420
 Db 361 YIARAGWTPGKYAWSILLDRSQTRLOIQLVISPFLIPVEDDVMERQRLIESVPDSVTP 420
 Qy 421 IIVEETTDIWINIHDFHVPQSHHEEIEFIPASECKTGFRHLKYITSLKESKYRSSG 480
 Db 421 IIVEETTDIWINIHDFHVPQSHHEEIEFIPASECKTGFRHLKYITSLKESKYRSSG 480


```

Db 707 ----- 706
Qy 781 GHPDQNEQGYLGVSAMQAEKPPSEPNRLLLHGFLENVHFAHTSILLSFLVRACKPYD 840
Db 707 ----- 706
Qy 841 LQYPOERHSIRVPESGEHVELHLLHYLQENLGSRIAALKVI 882
Db 707 --IYPOERHSIRVPESGEHVELHLLHYLQENLGSRIAALKVI 746

RESULT 13
ADT04044
ID ADT04044 standard; protein; 738 AA.
XX
AC ADT04044;
XX
DT 30-DEC-2004 (first entry)
XX
DE Human protein modification and maintenance molecule protein SeqID24.
XX
DE protein modification and maintenance molecule; PMMM; cytosstatic;
XX
KW immunomodulator; expression; immune disorder; cancer; human.
XX
OS Homo sapiens.
XX
PN WO2004084806-A2.
XX
PD 07-OCT-2004.
XX
PF 16-MAR-2004; 2004WO-US008006.
XX
PR 21-MAR-2003; 2003US-0456864P.
XX
PR 03-APR-2003; 2003US-0460512P.
XX
PR 13-MAY-2003; 2003US-0472027P.
XX
PR 30-MAY-2003; 2003US-0475072P.
XX
PR 30-MAY-2003; 2003US-0475190P.
XX
PR 24-JUN-2003; 2003US-0482575P.
XX
PR 13-AUG-2003; 2003US-0495151P.
XX
PA (INCY-) INCYTE CORP.
XX
PI Baughn MR, Marquis JP, Kable AE, Chawla NK, Emerling BM, Lee SY;
PI Hafalia AJA, Ramkumar J, Richardson TW, Wang JT, Nakamura LL;
PI Yang J, Jin P, Becha SD, Wilson AD;
XX
DR WPI; 2004-710256/69.
DR N-PSDB; ADT04072.
XX
PT New human protein modification and maintenance molecules (PMMM)
PT polypeptide, useful in preparing a composition for treating a disease
PT associated with decreased expression or overexpression of PMMM e.g.,
PT cancer.
XX
PS Claim 1; SEQ ID NO 24; 222pp; English.
XX
CC This invention relates to novel human protein modification and
CC maintenance molecule (PMMM) proteins and the DNA sequences which encode
CC them. The invention may be useful for the production of compounds with a
CC cytosstatic or immunomodulator activity. The proteins are useful in
CC preparing a composition for diagnosing or treating a disease or condition
CC associated with decreased expression or overexpression of functional
CC PMMM, for example immune disorders or cancer. The present sequence is
CC that of a human protein modification and maintenance molecule (PMMM)
CC protein of the invention.
XX
SQ Sequence 738 AA;

Query Match 81.2%; Score 3817; DB 8; Length 738;
Best Local Similarity 83.0%; Pred. No. 0;
Matches 732; Conservative 1; Mismatches 5; Indels 144; Gaps 3;
Qy 1 MAAAMETEQLGVIEFTADCEENIESQDRPKLEPPYVERYSWSQLKLLADTRKYHYMM 60

```

```

Db 1 MAAAMETEQLGVIEFTADCEENIESQDRPKLEPPYVERYSWSQLKLLADTRKYHYMM 60
Qy 61 AKAPHDFMFKRNDPDGPHSDRIYYLAWSGENRENTLFYSEIPKTIINRAA VMLSWKPLL 120
Db 61 AKAPHDFMFKRNDPDGPHSDRIYYLAWSGENRENTLFYSEIPKTIINRAA VMLSWKPLL 120
Qy 121 DLFOATLDYGMYSREBELLRERKRTGTGVIASDYHQSGCTFLQAGSGIYHVXGQPG 180
Db 121 DLFOATLDYGMYSREBELLRERKRTGTGVIASDYHQSGCTFLQAGSGIYHVXGQPG 180
Qy 181 FTQOPLRPNLVETSCPNIRMDPKLCPADPDWTAFTHSNDIWI SINIVTREERLLYVNNEL 240
Db 181 FTQOPLRPNLVETSCPNIRMDPKLCPADPDWTAFTHSNDIWI SINIVTREERLLYVNNEL 240
Qy 241 ANMEEDARSAGVATFVLOEEFDRYSGYMWCPKAEITPSGGKILRLIYEENDESEVEIHV 300
Db 241 ANMEEDARSAGVATFVLOEEFDRYSGYMWCPKAEITPSGGKILRLIYEENDESEVEIHV 300
Qy 301 TSPMLETRRADSPRYPKTGTANPKVTFKMEIMIDABGRIIDVIDKELIQFELFEGVE 360
Db 277 -----TANPKVTFKMEIMIDABGRIIDVIDKELIQFELFEGVE 317
Qy 361 YIARAGWTEGKYAMSIILLDRSOTRLQIVLISPFLIPVEDDVMERQRLIESVDSVTPL 420
Db 318 YIARAGWTEGKYAMSIILLDRSOTRLQIVLISPFLIPVEDDVMERQRLIESVDSVTPL 377
Qy 421 IYVEETTDIWINIHDFHVPFQSHBEEIEFIPASECKTGFRHLKYKITSILKESYKSSG 480
Db 378 IYVEETTDIWINIHDFHVPFQSHBEEIEFIPASECKTGFRHLKYKITSILKESYKSSG 437
Qy 481 GLPAPSDFKCPIKEEIAITSGEWEVLGRHGSNIQDVRRLVYFEGTKDSPLEHLLYVVS 540
Db 438 GLPAPSDFKCPIKEEIAITSGEWEVLGRHGSNIQDVRRLVYFEGTKDSPLEHLLYVVS 497
Qy 541 YVNPGEVTRLDRGYSHSCCISQHCDFPISKYSNOKNPHCVSLYKLSPPEDDPTCKTEF 600
Db 498 YVNPGEVTRLDRGYSHSCCISQHCDFPISKYSNOKNPHCVSLYKLSPPEDDPLA-NKEF 556
Qy 601 WATILDSAGPLPDYTPPEIFSFESTTGTLYGMLYKPHDLQPGKKYPTVFLFYGGPQVOL 660
Db 557 WATILDSAGPLPDYTPPEIFSFESTTGTLYGMLYKPHDLQPGKKYPTVFLFYGGPQ--- 613
Qy 661 VNNRFKGVKYFRLNTLASLGYVWIDNRGSCHRGLKPEGAFKYMKGQIEIDDDQVEGLQY 720
Db 614 ----- 613
Qy 721 LASRYDFIDLDRVGIHGSYGGVLSLMLMQRSDIFRVAIAGAPVTLWIFDGTGTYTERYM 780
Db 614 -----VAIAGAPVTLWIFDGTGTYTERYM 636
Qy 781 GHPDQNEQGYLGVSAMQAEKPPSEPNRLLLHGFLENVHFAHTSILLSFLVRACKPYD 840
Db 637 GHPDQNEQGYLGVSAMQAEKPPSEPNRLLLHGFLENVHFAHTSILLSFLVRACKPYD 696
Qy 841 LQYPOERHSIRVPESGEHVELHLLHYLQENLGSRIAALKVI 882
Db 697 LQYPOERHSIRVPESGEHVELHLLHYLQENLGSRIAALKVI 738

RESULT 14
ABB97362
ID ABB97362 standard; protein; 724 AA.
XX
AC ABB97362;
XX
DT 27-JUN-2002 (first entry)
XX
DE Novel human protein SEQ ID NO: 630.
XX
KW Human; antianemic; vulnerary; antinflammatory; immunomodulator;
KW antinfertility; cerebroprotective; cytostatic; rheumatic; gene therapy;
KW neuroprotective; antiparkinsonian; protein therapy; EST;

```


CC (DPRP). The dipeptidyl peptidase IV-related proteins (DPRP) and nucleic
CC acids encoding them are useful for treating infections such as fungal,
CC bacterial, protozoan and viral infections, particularly infections caused
CC by human immunodeficiency virus (HIV-1 or HIV-2), pain, diabetes,
CC precocious puberty, infertility, obesity, anorexia, bulimia, Parkinson's
CC disease, acute heart failure, hypotension, hypertension, urinary
CC retention, osteoporosis, angina pectoris, stroke, ulcers, asthma,
CC allergies, cancers, migraine, vomiting, psychotic and neurological
CC disorders (e.g. anxiety, dementia, or schizophrenia), and dyskinesias.
CC These may also be used in discovering therapeutic agents for the
CC treatment of reproductive, inflammatory and metabolic disorders. ABG61591
CC -ABG61612 represent human DPRP proteins
XX
XX

SQ		Sequence 658 AA;
Query Match		74.6%; Score 3504; DB 5; Length 658;
Best Local Similarity		100.0%; Pred. No. 0;
Matches 655; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	1	MAAAMETEQLGVEIFETADCEINTESQDRPKLEPFYVERYSWSQLKLLADTRKYHGYMM 60
Db	1	MAAAMETEQLGVEIFETADCEINTESQDRPKLEPFYVERYSWSQLKLLADTRKYHGYMM 60
QY	61	AKAPHDFMFKVENDPDGPHSDRIYYLAMSGENRENTLFYSEI PKTINRAAVLMLSWKPLL 120
Db	61	AKAPHDFMFKVENDPDGPHSDRIYYLAMSGENRENTLFYSEI PKTINRAAVLMLSWKPLL 120
QY	121	DLFOATLDYGMTSREBELLRERKRIGTGVGTASDYHOGSGTFLFOAGSGIYHVKGQPG 180
Db	121	DLFOATLDYGMTSREBELLRERKRIGTGVGTASDYHOGSGTFLFOAGSGIYHVKGQPG 180
QY	181	FTQOPLRNPLVETSCPNRMDPKLCPADPDWIAFIHSNDIWSNIVTREERLTYVHNEL 240
Db	181	FTQOPLRNPLVETSCPNRMDPKLCPADPDWIAFIHSNDIWSNIVTREERLTYVHNEL 240
QY	241	ANNEEDARSAGVATVLOEEFDYSGYWCPCPAETTPSGGKILRLIYEENDESEVEIHHV 300
Db	241	ANNEEDARSAGVATVLOEEFDYSGYWCPCPAETTPSGGKILRLIYEENDESEVEIHHV 300
QY	301	TSPMLETRRADSFYRKGTGTANPKVTFKMSIMIDAGRIIDVIDKELIOPPEILPEGVE 360
Db	301	TSPMLETRRADSFYRKGTGTANPKVTFKMSIMIDAGRIIDVIDKELIOPPEILPEGVE 360
QY	361	YIARAGWTEGKYANSILDRSQTQLQVLISPELFPVEDDVMERORLIESVPDSVTPL 420
Db	361	YIARAGWTEGKYANSILDRSQTQLQVLISPELFPVEDDVMERORLIESVPDSVTPL 420
QY	421	IYEEETDIWINIHDIHFVFPQSHEEIEFIFASECKTGFRHLYKITSILKESKYKSSG 480
Db	421	IYEEETDIWINIHDIHFVFPQSHEEIEFIFASECKTGFRHLYKITSILKESKYKSSG 480
QY	481	GLPAPSDFKCPKEIAITSGWEVLGRHGSNIQVDEVRRLVYPEGTKDSPLEHLYVVS 540
Db	481	GLPAPSDFKCPKEIAITSGWEVLGRHGSNIQVDEVRRLVYPEGTKDSPLEHLYVVS 540
QY	541	YVNPGEVTLTRDGRSHSCCISOHCDFFISKYSNOKNPHCVSLYKLSPPEDDPTCKTKEF 600
Db	541	YVNPGEVTLTRDGRSHSCCISOHCDFFISKYSNOKNPHCVSLYKLSPPEDDPTCKTKEF 600
QY	601	WATILDSAGPLPDYTPPEIFSPESITGFTLYGMLYKPHDLQPGKKYPTVLFIYGG 655
Db	601	WATILDSAGPLPDYTPPEIFSPESITGFTLYGMLYKPHDLQPGKKYPTVLFIYGG 655

Search completed: April 14, 2006, 12:46:02
Job time : 199 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 14, 2006, 12:46:23 ; Search time 46 Seconds
(without alignments)
1844.852 Million cell updates/sec

Title: US-10-825-632-1
Perfect score: 4700
Sequence: 1 MAAAMTEQLGVIEFTADC.....HLHYLOENLGSRIAALKVI 882

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	922.5	19.6	931	T32919	hypothetical prote
2	842	17.9	738	A87516	dipeptidyl peptida
3	754	16.0	741	JC5142	X-Pro dipeptidyl-p
4	620	13.2	711	S66261	X-Pro dipeptidyl-p
5	529	11.3	766	CDHU26	dipeptidyl-peptida
6	519.5	11.1	793	T41703	dipeptidyl aminope
7	518.5	11.0	760	S23752	dipeptidyl-peptida
8	510.5	10.9	792	A39914	dipeptidyl-peptida
9	484.5	10.3	818	A30107	dipeptidyl aminope
10	465	9.9	931	A49737	dipeptidyl aminope
11	451.5	9.6	799	T25174	hypothetical prote
12	445.5	9.5	779	T25173	dipeptidyl aminope
13	440	9.4	803	A41793	dipeptidyl aminope
14	427	9.1	803	I68600	dipeptidyl aminope
15	427	9.1	865	T154331	hypothetical prote
16	378	8.0	829	T19514	dipeptidyl-peptida
17	374	8.0	795	F82858	dipeptidyl-peptida
18	331.5	7.1	759	I38593	fibroblast activat
19	293	6.2	743	T37700	probable dipeptid
20	282	6.0	642	C71137	hypothetical prote
21	269.5	5.7	631	T875007	probable acylamino
22	268	5.7	683	B87495	prolyl oligopeptid
23	255.5	5.4	709	B82580	alanyl dipeptidyl
24	243.5	5.3	657	E70025	probable acylamino
25	236	5.0	622	F71174	hypothetical prote
26	223.5	4.8	632	E75057	peptidase PAB1418
27	217.5	4.6	536	F90299	acylaminoacyl-pept
28	200.5	4.3	591	H72474	probable acylamino
29	185	3.9	569	S74053	probable acylamino

30	183.5	3.9	608	2	F83397	probable peptidase
31	182.5	3.9	674	2	B84381	acylaminoacyl-pept
32	180.5	3.8	721	2	T09631	probable acylamino
33	180	3.8	667	2	A87711	prolyl oligopeptid
34	172	3.7	732	1	JC4655	acylaminoacyl-pept
35	171	3.6	732	1	S07624	acylaminoacyl-pept
36	170.5	3.6	572	2	F72455	probable acylamino
37	163.5	3.5	659	2	F72568	probable acylamino
38	161	3.4	676	2	C97775	acylamino-acid-rel
39	160.5	3.4	732	1	JU0132	acylaminoacyl-pept
40	160.5	3.4	745	2	T33751	hypothetical prote
41	158	3.4	598	2	F84199	hypothetical prote
42	157	3.3	614	2	E75094	prolyl endopeptida
43	155.5	3.3	654	2	AD3183	peptidase [importe
44	155	3.3	629	2	T15945	hypothetical prote
45	154.5	3.3	828	2	G87584	hypothetical prote

ALIGNMENTS

RESULT 1

T32919

hypothetical protein K02F2.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004

C:Accession: T32919

R:Maggi, L.; Goela, D.

submitted to the EMBL Data Library, January 1998

A:Description: The sequence of C. elegans cosmid K02F2.

A:Reference number: Z21246

A:Accession: T32919

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-931 <MAG>

A:Cross-references: UNIPROT:O44987; UNIPARC:UPI0000076BD8; EMBL:AF043699; PIDN:AAB97564

A:Experimental source: strain Bristol N2; clone K02F2

C:Genetics:

A:Gene: CBSP:K02F2.1

A:Map position: 1

A:Introns: 58/3; 82/2; 131/2; 178/2; 275/3; 322/2; 404/2; 441/1; 464/2; 486/3; 528/2; 5

Query Match	19.6%	Score	922.5	DB	2	Length	931
Best Local Similarity	29.3%	Pred. No.	2.1e-57				
Matches	282	Conservative	145	Mismatches	326	Indels	211
Gaps	41						
QY	33	EPFVYRYSQKLLADTRKYHYGMKAPHDHFMVKRNDPDGPHSDRIYVYLAMSGEN	92				
DB	36	EPARFETRSFSQ---LIDHARSWKTEVRGTTGGFTKISLMRAE---KDRLNMYAISVSP	89				
QY	93	RENT--LPYSEIP-KTINRAAVLMSWKPLLDLFQATLDYGMV-----SRE	135				
DB	90	GTNTQSFSVTIPLEVEKAQVADRK-----FELKLSGYNVDYIRMSCKRTPPSAE	142				
QY	136	EELLRKRKIGTV-GIASYDHYOG-----SGTFLPQAG-----	167				
DB	143	FTLOCQRORSQVVTGIDYEIRNGKMLMAGDQLFRYNPLNEALAAPIAVPDDQSSTEP	202				
QY	168	-----SGYHVKGQPGQ-----FTQOQLRPNLVETSCP-----NIRMPKLC	206				
DB	203	MDISEGTSITSGKCSNEAFOSSTVPVPTRIPRIKKTTTEKPTATPTNTNFVSSAKVCP	262				
QY	207	ADPOWIAFIHSDNIWISNIVTREERLTVVHNELANMEEDARSAGVATFVLQEEFDYRS	266				
DB	263	ADSSLLAVLNKQVI-----EKNGKIIHRTSSNSKHI TN--GVPSYIVQEELEPER	313				
QY	267	YWCFCFAETTPSGKILRIIYENDSEVEI IHV-----TSPMLETRRADSFYRPTKG	319				
DB	314	IWW-SESKT-----RLLYEHVNEKVAESQGVNGDPPVAPM-----KYPRAG	355				
QY	320	TANPKVTFNGEIMIDAGRIIDVIDKELIQPEILFEG---VEYIARAGWTPEGKYAWS	376				
DB	356	TKNAYSTLRM---VILENGKAYDVPLKD-----EVITYKHCFFYEYITRAGFFSDGTTVMV	407				

QY	377	ILLDRSQRLOQLVLI-----SPELFIPVEDDMERQRIIESVP	414
Db	408	QVMSRDOAQCSLLIIPYDFLLPBELOGSKEONLQLSTDLNNGVWDD--KSHEETMEXPP	466
QY	415	DSVT--PLIIVETTIDMINIHDIHFVFPQSHBEE--IEFIFASECKTGFRHLKYITSIL	470
Db	467	RGKLRGTQIHKARNDYWINTHNAIYPLKIDDEHPMYEPIYCLEPNG--SCLALISABEL	525
QY	471	KESKYKRSSGGLPAPSDFKCPKKEEIAITSGEWEVLGRHGSNIQVDEVRRLVYFEGTKDS	530
Db	526	DQNGY-----CRHTEKLLMAENFSI--NKSNGIIVDVEVRELVYVANESH	569
QY	531	PLEHHLVYVYNPGEVTRLTDRGYSHSCCISOHCDFEISKYSNOK-----NPHCVSLYKL	586
Db	570	PTEWNI--CVSHRTQHQALTESGI-----C--FKSERANGKALDLDHGFACYMT	617
QY	587	S--SPEDDPTCKTKEP-W-----ATILDSAGP-LPD--YTPPEIFSFES--TTG	627
Db	618	SVGSPAE---CRPYGFRWKENEVLPSTVAAITVSHGQOPDLHFDSPEMIEFQSKTG	674
QY	628	FTLYGMLYKPHDLQPGKKYPTVLFIYGGPQVOLVNNRPFKGVKYFRLNTLASLGYVVVVID	687
Db	675	LMHYAMILRPSNFPDKYKYPVHYVYGGG-IQIHNDFSWIQYIR---FCRLGYVVVVFI	731
QY	688	NRGSHRGKLFEGAFKYNQGTIEDDQVEGLQYLASRY-DFIDLDRVGIHGWISYGYLSL	746
Db	732	NRGSAHRGIEFERHIHKMGTVVEVDQVEGLQMLAERTGGFMDSRVVHGWISYGYMAL	791
QY	747	MALMQRSDIFRVAIAGAPVTLMIFYDTGYTERYMGHPDQNEQGYLGSVAMQAEPSPSEP	806
Db	792	QMLAKHPNIYRAAIAAGAVSDWRLYDTATYTERYMGYP--LEEHVYGAISSITGLVEKLDPDEP	850
QY	807	NRLLLHGLFDENVHFAHTSILLSPFLVRAGKPYDLOIYYPQERHSIRVPESGEHYELHLH	866
Db	851	NRLMLVHGLMDENVHFAHTLHLVDECIKKQWHELVIFFPNRHHGVRNNDASIYLDARMY	910
QY	867	YLOE	870
Db	911	FAQQ	914
RESULT 2			
A:Species: Caulobacter crescentus			
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004			
A:Accession: A87516			
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J. B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M. Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001			
A>Title: Complete Genome Sequence of Caulobacter crescentus.			
A:Reference number: A87249; MUID:21173698; PMID:11259647			
A:Accession: A87516			
A>Status: preliminary			
A:Molecule type: DNA			
A:Residues: 1-738 <STO>			
A:Cross-references: UNIPROT:Q9A6E0; UNIPARC:UPI00000C7616; GB:AE005673; NID:913423647; F			
A:Gene: CC2154			
Query Match			
Best Local Similarity 17.9%; Score 842; DB 2; Length 738;			
Matches 221; Conservative 115; Mismatches 262; Indels 141; Gaps 24;			
QY	133	SREELARERKRIGTIGIASYDVHOGSGTFLFOAGSGIY--HVXKGGPGQFTQQLRPNL	190
Db	94	SEAEKARRERARVSARGIVEYSWDRQGRFILVPLDGLYLDADAADGKITRUTE-----	146
QY	191	VETSCPNIWMDPKLPADPDWIAFIHSDNIWISNIVTREERLITVYHNELANMEEDARSA	250
Db	147	-----TPGDEVDKVPKG--GVSVYVRDQNLIKPVAGGAETALT-----TDGKDLSF	194

QY	251	GVATFVLOEBEDRSGYVWCPKAAETTPSGGKILRLIYLEENDESEVEIHHVTSMLETRRA	310
Db	195	GVAEFIVORELDRFRTGYWSPDES-----RIVYTRVDESGVIV-----PRA	236
QY	311	D-----SFRPQTCTANPKVTFKMSIMIDAEGRI--IDV-TDKELIQPPPELLPSG	358
Db	237	DIGPGGATVWQRYPRAGRPNNAVVDLFRDL---ASGKVTAOLDLGANKOI-----	283
QY	359	VEYIARAGWTPEGKYAMSILLDRSQTRLQIVLISPELFIPEVDDVMERQRLIESVDSTV	418
Db	284	--YVARVAMSADGKTYYVQRLSRDQKTLDLLAF-----DAAT	318
QY	419	PL--IYIETTDIWINIHDIHFVFPQSHBEEIEFIFASECKTGFRHLKYITSILKESYK	476
Db	319	GAGKTILATDHPHFEVSNDRFRLTDG-----TFLMGSE-KDGNQHLRYA-----	363
QY	477	RSGGLPAPSDFKCPKKEEIAITSGEWEVLGRHGSNIQVDEVRRLVYFEGTKDSPLEHHL	536
Db	364	-ADGKLI-----QITKGDWPIVIGLE-----VDEARKVAIFASIDTPIERRL	406
QY	537	YVSYVYNPGEVTRLTDRGYSHSCCISOHCDFEISKYSNQNPKHCVSLYKLSSP-----	591
Db	407	YEVSYAKGPKKALTSAGGWAAKVADNGGAFAGTYSDPKTPSQTALYSADGKRVRWIEE	466
QY	592	DPCTCKTEFWATILDSAGPLDPTTPEIFSPESSTGTFLYMLYKPHDLQPGKKYPTVL	651
Db	467	NKLAEGHPY---PYAANLPQ---PEFGSLKAADGETLHVEILKPIGDFPAKKYPAIVS	519
QY	652	IYGGPQVOLVNNRPFKGVKYFRLNTLASLGYVVVVVDNRGSHRGKLFEGAFKYMGOIEI	711
Db	520	YTGPGHAQRYMKNWHPSE---RTYLEAGYVIFKLDNRGSGNRSKAPKMRALDRKLTVEV	576
QY	712	DDQVEGLQYLASRYDFIDLDRVGIHGWISYGYLSIMALMQRSDIFRVAIAGAPVTLWIF	771
Db	577	EDQLLGAKEFLAQ--FYVDADKLGVMWISYGGFMALMLLTAEPTPFKAGAAGAPTESL	635
QY	772	DTGYTERYMGHPDQNEQGYLGSVAMQAEPSPENRLLHGLFDENVHFAHTSILLSP	831
Db	636	DTATYTERYMGKPDENKAGAYSDINNRIKUL--APGSLLLHGMADONVIFENSTRLMAA	693
QY	832	LVRAGKPYDLOIYYPQERHS	850
Db	694	LQRKAILFEMAMYPGERHS	712

RESULT 3

JC5142

X-Pro dipeptidyl-peptidase (EC 3.4.14.11) - Xanthomonas maltophilia
C:Species: Xanthomonas maltophilia
C:Date: 31-Jan-1997 #sequence_revision 27-Feb-1997 #text_change 09-Jul-2004
C:Accession: JC5142
R:Kabaehima, T.; Ito, K.; Yoshimoto, T.
X:Biochem, 120, 1111-1117, 1996
A>Title: Dipeptidyl peptidase IV from Xanthomonas maltophilia: Sequencing and expression
A:Reference number: JC5142; MUID:97164011; PMID:9010758
A:Accession: JC5142
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-741 <KAB>
A:Cross-references: UNIPROT:P95782; UNIPARC:UPI00000B5F8A; DDBJ:D83263; NID:gi1753196; P
C:Comment: This enzyme catalyzes the hydrolytic removal of amino-terminal dipeptidyl re
C:Keywords: dipeptidyl-peptidase IV
F:4-18/Domain: transmembrane #status predicted <TM>
F:610/Active site: Ser #status predicted
F:685/Active site: Asp #status predicted
F:717/Active site: His #status predicted

Query Match 16.0%; Score 754; DB 2; Length 741;
Best Local Similarity 28.4%; Pred. No. 1.5e-45;
Matches 214; Conservative 120; Mismatches 281; Indels 138; Gaps 26;

133 SREELLRRKRIGTV-GIASYDHYHGGSGTFLFOAGSG--IYHVKGQGGTQOQPLRN 189
134 SDEKARRRQRIAAWTGVDYQSPDAQRLLFLPGGELYLDLQEGKAARQL-----147
190 LVETSCPNIRMDPKLCPADPDWAFIHSNDIWIHNIIVTRERRLLTYVHNELANWEEDARS 249
148 ---THGEGFATDKLSPKG-GFVSFIRGNLWIDLASGRQMOLT-----ADGST 193
250 A---GVATFVLOEEDRYSGYWCPCAEETTPSGKILRLIYEENDESEV-----EIIHV 300
194 TINGGAEFAVEADEMRHTGYWAPDASA-----IAYARIDESPVVQKREYEVAD 244
301 TSPMLRETRADSRYPKPTGTANPKVTFKMSIMIDAEGRIDVIDKELIOPFEILPEGVE 360
245 RTDVIQ-----RYPAAGDANVQKLGWISPAEQATQWIDLKEQDI-----287
361 YIARAGWTPEGKAVMSILLDRSQTRQLQIVLISPELFIPEVDDVNEROLIESVPDSVTP 420
288 YLARVNVWRDPQLSFO-RQSDQKGLDLVEVTL-----ASNQOR-----325
421 ILYEETTDIWINIHDIHFVPPQSHHEEIEFIPASECKTGPRHLYKITSILKESKYKRSG 480
326 VLAHETSPTWPLHNSLREL-----DQGSILWSSB-RIGFOHLYRI-----DSGKAA--372
481 GLPAPGDFKCPKEETAITSGEWEVLGRHSGNIQVDEVRRLVYPEGTKDSPLEHLYVVS 540
373 -----ALTHGNWSV-----DELLAVDEKAGLAYFRAGIESARESQIYAVP 412
541 YVNPGEVTRLDRGYSHSCCISQHCDFPISKYKQKQPHCVSLYSKQKPHCVSLYKL 586
413 -LQGGQPORLSKAPGMSHASFARNASVVDNSNNSTPQTELFRANG-----EK 461
601 WAIL--DSAGPLPDYT-----PPEIFSESTTGT-LYGMLYKPHDLQPKKYPTVL 650
462 IATLVENDLADKHPYARYEARQRPVEFGTLTAADGKPLNYSVIKPAGDPDAKRYPAV 521
651 FYGGQVQVNNRFGVKYFRLNT--LASLGVVVVVVVDNRGSGCHRGKLFEGAFKYMGOI 709
522 YVGGPASQTVTDSWFERGDHFLFNQYLAQGVVVFSLDNRGTGPRGRDFGALYKQGT 581
710 EIDQVEGLQYLASRYDFIDLRVGHGWSYGGYLSLALMQRSDIFRVAIAGAPVTLWI 769
582 EVADQLRGVAVL--KQCPWVDPAIRIGVQGSNGSYMTLLAKASDSYACGVAGAPVTDWG 640
770 FYDTGYTERYMGHPDQNEQGYLGSVAMQAEKFPSEPNRLLLHGFEDENVHAFTSILL 829
641 LYDSHYTERYMDLPARDNAGYREARVLTHIEGLRSP---LLLIHGMADDNVLFTNSTLM 697
830 SFLVRACKPYDLQIYQERHSIRVPESGHEYL 862
698 SALQKRGQPELMTYPGAKHGSLGADALHRYRV 730

RESULT 4

S66261
X-Pro dipeptidyl-peptidase (EC 3.4.14.11) - Flavobacterium meningosepticum
C:Species: Flavobacterium meningosepticum
C>Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C:Accession: S66261
R:Kobashima, T.; Yoshida, T.; Ito, K.; Yoshimoto, T.
Arch. Biochem. Biophys. 320, 123-128, 1995
A>Title: Cloning, sequencing, and expression of the dipeptidyl peptidase IV gene from *P.*
A:Reference number: S66261; MUID:95314307; PMID:7793970
A:Accession: S66261
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-711 <KAB>
A:Cross-references: UNIPROT:Q47900; UNIPARC:UPI00000B2115; EMBL:D42121; NID:G577283; PII
A:Superfamily: dipeptidyl-peptidase IV
C:Keywords: dipeptidylpeptide hydrolase

Query Match 13.28; Score 620; DB 2; Length 711;
Best Local Similarity 25.28; Pred. No. 5.2e-36;

Matches 195; Conservative 134; Mismatches 284; Indels 160; Gaps 26;
145 IGTGVIASDYDHYG-----SGTF-----LFQAGS-----GIYHVKG 176
54 IEPTGIKYSYKTSQKEKNIVDGSGFQYTFNSDESKILLQKSSQSIYRHSFLGFEVKD- 112
177 GPOGTQOQLRP-NUVETSCPNIRMDPKLCPADPDWAFIHSNDIWIHNIIVTRERRLLTY 235
113 -----LKSRVTVSLNNAWIOEPKPSF-DGSKVAFIADNNLFYODLNTGKITQIT 162
236 --VHNELANWEEDARSAGVATFVLOEEDRYSGYWCPCAEETTPSGKILRLIYEENDES 293
163 DGKKNELIN-----GLGDWYEEFGHADIYQ-----NKAGDALVFVREDKVP 208
294 EVELI-IHVTS--PMLRETRADSRYPKPTGTANPKVTFKMSIMIDAEGRIDVIDKELI 350
209 EINIPYQNLVPELMT-----VKYPKAGEANSATVAYLYQL-----SSGK-----249
351 PFEILPEGVE--YIARAGWTPEGKAVMSILLDRSQTRQLQIVLISPELFIPEVDDVNER 408
250 SAQLNFGSSEKYYIPQLFOTNANDEIVVATANRHNQKVDLLKVNTK-----295
409 LIESVPDSVTPPLIYBETTDIWINIHDIHFVPPQSHHEEIEFIPASECKTGPRHLYKITS 468
296 -----TAAVSKLFTETONAMETDNLMEF-----LDNSFLWASE-RDGRHLYWYDA 343
469 ILKESKYKRSGGLPAPSDFKCPKEETAITSGEWEVLGRHSGNIQVDEVRRLVYFEGTK 528
344 AGKLKK-----QVSKGDWEIINYYGYNPKTKE-----VVIQTTE 377
529 DSPLEHLYVVS--YVNPGEVTRLDRGYSHSCCISQHCDFPISKYKQKQPHCVSLYKL 586
378 KGSINK--VVSKLINTGKTQLLSNAEGNSAAFSKTFNYFINTSSTAKVPTKYLKDA 434
587 SSPE-----DDPTCKTEFWATILDSAGPLPDYTPPEIFSESTTGTFLYGMLYKPHD 639
435 NGKDVKELQNNDLLNKLKS-----DNFAKEFIIIPNAGDQWANNWIKPN 482
640 LQPKKYPTVLFIYGGPQVQVNNRFGVKYFRLNTLASLGVVVVVVVDNRGSGCHRGKLF 699
483 FDPAKKYPVFMFYSGSGSQVANSWDGNGIWFMDMLAQGYLVVVCVDRGTGFRGTKYK 542
700 GAFKYMGOI EIDQVEGLQYLASRYDFIDLRVGHGWSYGGYLSLALMQRSDIFRVA 759
543 KVTYNKLGKYEIEDQITAAKWLGNQ--SYVDKSRIGIFGWSYGYMASLMTKGDVFPKMG 601
760 IAGAPVTLWIFDTGYTERYMGHPDQNEQGYLGSVAMQAEKFPSEPNRLLLHGFEDEN 819
602 IAVAPVNWRFYDSIYTERFLQTPQENKQDGLNPSPTTAKLL---KGKFLIHGTADDN 658
820 VHPAHTSILLSFLVRACKPYDLQIYQERHSIRVPESGHEYLHLLHLYQENL 872
VHFQNSMEFSEALIQNKQDFMAYPDKNHSIIGGNTRPQLYEKMTNILENL 711

RESULT 5

CDHU26
dipeptidyl-peptidase IV (EC 3.4.14.5) - human
N:Alternate names: cell surface glycoprotein CD26; thymocyte-activating molecule (THAM)
C:Species: Homo sapiens (man)
C>Date: 31-Dec-1993 #sequence_revision 23-Aug-1996 #text_change 09-Jul-2004
C:Accession: S24313; B42408; B61136; S59510; I56154; S59857; S15520
R:Misumi, Y.; Hayashi, Y.; Arakawa, F.; Ikehara, Y.
Biochim. Biophys. Acta 1131, 333-336, 1992
A>Title: Molecular cloning and sequence analysis of human dipeptidyl peptidase IV, a se-
A:Reference number: S24313; MUID:92329551; PMID:1352704
A:Accession: S24313
A:Molecule type: mRNA
A:Residues: 1-6, 'I', 8-766 <MIS>
A:Cross-references: UNIPROT:P27487; UNIPARC:UPI00001GAE80; EMBL:X6708; NID:G35335; PII
R:Darmoul, D.; Lacasa, M.; Baricault, L.; Marguet, D.; Sapin, C.; Trotot, P.; Barbat, A.
J. Biol. Chem. 267, 4824-4833, 1992
A>Title: Dipeptidyl peptidase IV (CD 26) gene expression in enterocyte-like colon cancer

Db 277 IKYKPGNPNPSVSLFVADLNSSNFSLMHNEPLAEP-----VQNVLMVNTS----- 326
QY 373 YAWSTLL---DRSOTRLQIVLISPEL--FIPVEDDVMER-----ORLIESVPDSVPTLLI 422
Db 327 ---SVLVQFTNRNSTCITARLLDTLTKSIHTVKTCLBEGWVEVOOSAKMFP--LNNSLV 381
QY 423 YEETDWINIHDFHVPFQSHBEEIEFTASECKTGFRHLKYKTSILKESKYRRSGGL 482
Db 382 WENWSD---GYFDLALDDYNNHLAFIPF-----NGSS 410
QY 483 PAPDFKCPKEETAITSGEWEVLGRHSGNIQVDEVRRLVYFEGTKDSPLEHLVYVYV 542
Db 411 P-----IYLTSGAWDVT---DGPIDHGDGFGNVYPLATLKDSTERHLYYS-L 454
QY 543 NPGFVRLTRDRGYS---HSCCISQHCDFFTSKYSNOKNPHCVSLYKLSPPDDPTCKTKE 599
Db 455 DTLEYGITDNGDEGYSTSFSPGDFYVLNVHGDVP---WQELASTKDQCYLSLE 510
QY 600 FWATILDSAGPLPYTPPEI-----FSPBSTTGFTLYGMYLKPDLQPKKYPVTLFYGG 655
Db 511 TNSRLKQO---LSSITLPSVEYKGLTFNDTT-FNF--MERRPRNFDVNVKYPVLFYAYGG 564
QY 656 PQVQLVNNRFKGVKYFRLNTLASLG-----YVVVIDNRGSGHRLKPEGAFYKMGQI 709
Db 565 PGGQOV-----AKLFRVDYFOAYLASHDPDEFIVTLVLDGRGTGFGNAPRYSVSRHLGEW 618
QY 710 EIDDOVEGLOYLASRGYDFIDLDRVGIHGWISYGGYLSMALMORSDFRVAIAGAPVTLWI 769
Db 619 ESYDQOQAGKFWAD-LFVDENHVGWISYGGYLTKLTL-ETQDVSFGYNAVAVTDWR 676
QY 770 FYDTGYTERYMGHPDQEOGYLGSVAMQAEKFPSEPNRLLHLLHVLQENLGSRIAAL 879
Db 677 LYDSVYTERYMDLPQYNKEG-YKNSQIHDYKFP-KQLKRFVFAHGTGDDNVHFQHSNHL 734
QY 830 SFLVRAG-KPYDLOIYQOEHSIRVPSGEHVELHLLHVLQENLGSRIAAL 879
Db 735 DGLNLANCYNVDMVFPDPSAHSI-----SYHNASLSYHRLSEWIGDALGRI 781

RESULT 7

S23752
diptidyl-peptidase IV (EC 3.4.14.5) alpha chain - mouse
N:Alternate names: CD26 alpha subunit; THAM alpha subunit
C:Species: Mus musculus (house mouse)
C>Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: S23752; A46465; A56030
R:Marguet, D.; Bernard, A.M.; Vivier, I.; Darmoul, D.; Naquet, P.; Pierres, M.
J. Biol. Chem. 267, 2200-2208, 1992
A:Title: cDNA cloning for mouse thymocyte-activating molecule. A multifunctional ecto-di
A:Reference number: S23752; MUID:92129288; PMID:1370813
A:Accession: S23752
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-760 <MAR>
A:Cross-references: UNIPARC:UPI0000172A2D; EMBL:X58384
R:Vivier, I.; Marguet, D.; Naquet, P.; Bonicel, J.; Black, D.; Li, C.X.; Bernard, A.M.;
J. Immunol. 147, 447-454, 1991
A:Title: Evidence that thymocyte-activating molecule is mouse CD26 (diptidyl peptidase
A:Reference number: A46465; MUID:91302787; PMID:1712807
A:Accession: A46465
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-20 <VIV>
A:Cross-references: UNIPARC:UPI0000172A2E
A:Experimental source: M14.T thymoma cells, Swiss nu/nu
A>Note: sequence extracted from NCBI backbone (NCBIP:42236)
R:Bernard, A.M.; Mattei, M.G.; Pierres, M.; Marguet, D.
Biochemistry 33, 15204-15214, 1994
A:Title: Structure of the mouse diptidyl peptidase IV (CD26) gene.
A:Reference number: A56030; MUID:95092780; PMID:7999781
A:Accession: A56030
A>Status: preliminary
A:Molecule type: DNA

A:Residues: 746-760 <BER>
A:Cross-references: UNIPARC:UPI0000172A2F; GB:U12620

C:Genetics:

A:Gene: CD26

C:Superfamily: diptidyl-peptidase IV

C:Keywords: diptidyl-peptide hydrolase; glycoprotein; transmembrane protein
F:213,223,315,514,679/Binding site: carboxylate (Asn) #status predicted
F:624,702,734/Active site: Ser, Asp, His #status predicted

Query Match 11.0%; Score 518.5; DB 1; Length 760;

Best Local Similarity 25.0%; Pred. No. 1e-28;

Matches 203; Conservative 107; Mismatches 306; Indels 197; Gaps 37;

QY 131 MYREELLRRKRIGTVGI-----ASYDHQSGGTLPF-----QAGS 168

Db 67 LYKQENILLNLAHGNSSIFLENSTFSGVHSVSPDLFLVLLYNYVKKWRHSYTSY 126

QY 169 GYHVKDGQGGFQQPLRENVLVETSCPNIRMDPKLCPADPDWI-----AFIHSND 219

Db 127 NIYDNK-----ROLITEKIPN-----NTQWITWSPGCHKLATVWKNQ 165

QY 220 IWNISNIVTRBERRITYVHNELANMBDARSAGVATFVLOEE-FDRISYGVWCPK----- 272

Db 166 IYKVEPHLSHRIT-----STGENVIYNGITDVIYEEVFGAYSALWSPNNTPLAY 219

QY 273 AETTPSGKILRLIYENDESEVEIIHVTSPMLETRADSFRYPKTGTANPKVTFKMEI 332

Db 220 AQFNDTGVPLEIYSF-YSDS-----LQYPKTWIPIPKAGAVNPTVKPFI--V 265

QY 333 MIDAEGLIIDVIDKELQPEILFEGVEIARAGWTPECKYA-----WSILLDRSQ 383

Db 266 NIDSLSSSSAAPIQIPAPASVA-RGDHYLQDVWATEERISLQWLRIONYSNAICDY 324

QY 384 TRLOVLISPELFPVEDDVMQRLESVPDSVTPLLIYEETDIWINIHDFHVPFQS 443

Db 325 DKINLTWNCPS-----EQQHV-----ENSTTGWG--RFRPAEHP 358

QY 444 HEEIEFIFASECKTFRHLKYKITSILKESKYRSGGLPAPSDPK-CPKEEIIATSGE 502

Db 359 TSDGSSFYKIISDKQYKICHF-----PKDKDCCTF-----ITKGA 395

QY 503 WEVLGRHSGNIQVDEVRRLVYFPGT-KDSPLEHLVYVSVNPNVGEVTRLTRDGYSHSCCI 561

Db 396 WEVI-----STEALTSYLYISNQYKEMPGGGRNLYKI-----QLTD--HTNVKCL 439

QY 562 S-----QHCDFISKYSNKNPH---C-----VSLYKLSPPEDDPTCKTKEFWATILDSAG 609

Db 440 SCDLNPERCQYVAVSFKEAKYQLQGWGPGPLPLYTLHRSTHKLRLVLE-----DNSA 493

QY 610 ---PLPDVTPPE---IFSFSSTTGFTLYGMYLKPDLQPKKYPTVLFYGGPQVQLVNN 663

Db 494 LDRMLQDVQMPKSKLDIFVLNETRF-WYQMLPPH-FDKSKYPLLLDLYAGPSCQ---- 547

QY 664 RPKGVKVPRLN---TLASLGYYVVV-IDNRGSGHRLKPEGAFYKMGQIEIDDOVEGLQ 719

Db 548 ---KADASFLNWTATLASTENIIVASFDGRGSGYQGDKIMHAINRRLGLEVEIDQEAR 605

QY 720 VLASRYDFIDLDRVGIHGWISYGGYLSMALMORSDFRVAIAGAPVTLWIFDVTGTYERY 779

Db 606 QFV-KMGFVDSKRVAVIWNYSYGGYVTSWLGSSGVFKGIAVAPVSRWEYDVSYTERY 664

QY 780 MG---HPDQEOGYLGSVAMQAEKFPSEPNRLLHLLHVLQENLHFAHTSILLSFLVRAGK 837

Db 665 MGLPIPEDNLHRYNSTVMSRAEHF--KQVEYLLIHGTADDNVHVFQSSAQISKVLVDAGV 722

QY 838 PYDLOIYQOEHSIRVPSGEHVELHLLHVLQ 870

Db 723 DFQAWMYTDEDHGIASSTAHQHYSHMSHFLQ 755

RESULT 8

A39914

diptidyl-peptidase IV (EC 3.4.14.5), membrane-bound form precursor - rat

N;Alternate names: GP110; membrane glycoprotein 110K; OX-61
N;Contains: dipeptidyl-peptidase IV, soluble form
C;Species: Rattus norvegicus (Norway rat)
C;Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text change 09-Jul-2004
C;Accession: A39914; A33315; A60730; A42203; S38949; A31781
R;Hong, W.; Doyle, D.
Proc. Natl. Acad. Sci. U.S.A. 84, 7962-7966, 1987
A;Title: cDNA cloning for a bile canalicular domain-specific membrane glycoprotein of rat
A;Reference number: A39914; MUID:88068516; PMID:3479775
A;Accession: A39914
A;Molecule type: mRNA
A;Residues: 1-792 <H0>
A;Cross-references: UNIPROT:P14740; UNIPARC:UPI000017098A; GB:J02997; NID:G204463; PIDN:
R;Ogata, S.; Misumi, Y.; Ikehara, Y.
J. Biol. Chem. 264, 3536-3604, 1989
A;Title: Primary structure of rat liver dipeptidyl peptidase IV deduced from its cDNA an
A;Reference number: A33315; MUID:89123496; PMID:2563382
A;Accession: A33315
A;Molecule type: mRNA
A;Residues: 1-37, 'A', 39-182, 'I', 184-331, 'T', 333-351, 'C', 353-393, 'V', 395-561, 'L', 563-623,
A;Cross-references: UNIPARC:UPI0000129841; GB:J04591; NID:G403973; PIDN:AAA41096.1; PID:
A;Note: the authors translated the codon GCG for residue 38 as Arg, ACC for residue 332
A;Accession: B33315
A;Molecule type: protein
A;Residues: 1-20,35-54,427-443;505-509;511-520;530-538;593-600;602-608;618-627 <H02>
A;Cross-references: UNIPARC:UPI0000172A2E; UNIPARC:UPI0000172A30; UNIPARC:UPI0000172A31;
A36; UNIPARC:UPI0000172A37
R;McCaughan, G.W.; Wickson, J.E.; Creswick, P.F.; Gorrell, M.D.
Hepatology 11, 534-544, 1990
A;Title: Identification of the bile canalicular cell surface molecule GP110 as the ectop
A;Reference number: A60730; MUID:90228896; PMID:1970322
A;Accession: A60730
A;Molecule type: protein
A;Residues: 28-47, 'XX', 50-53,55-58 <MCC>
A;Cross-references: UNIPARC:UPI0000172A38
R;Ogata, S.; Misumi, Y.; Teuji, E.; Takami, N.; Oda, K.; Ikehara, Y.
Biochemistry 31, 2582-2587, 1992
A;Title: Identification of the active site residues in dipeptidyl peptidase IV by affini
A;Reference number: A42203; MUID:92190188; PMID:1347701
A;Accession: A42203
A;Molecule type: Protein
A;Residues: 'R', 625-630, 'X', 632-648 <OC2>
A;Cross-references: UNIPARC:UPI0000172A39
R;Iwaki-Egawa, S.; Watanabe, Y.; Fujimoto, Y.
Biol. Chem. Hoppe-Seyler 374, 973-975, 1993
A;Title: N-terminal amino acid sequence of the 60-kDa protein of rat kidney dipeptidyl p
A;Reference number: S38949; MUID:94128239; PMID:7903271
A;Accession: S38949
A;Status: preliminary
A;Molecule type: protein
A;Residues: 281-302 <IWA>
A;Cross-references: UNIPARC:UPI0000172A3A
R;Hong, W.; Doyle, D.
J. Biol. Chem. 263, 16892-16898, 1988
A;Title: Membrane orientation of rat gp110 as studied by in vitro translation.
A;Reference number: A31781; MUID:89034185; PMID:3182821
A;Accession: A31781
A;Molecule type: mRNA
A;Residues: 1-40 <H03>
A;Cross-references: UNIPARC:UPI0000172A3B
C;Comment: This protein is localized to the bile canalicular, which is the apical domain
C;Superfamily: dipeptidyl-peptidase IV
C;Keywords: dipeptidyl-peptidase; glycoprotein; homodimer; liver; serine proteina
F;1-28/Product: dipeptidyl-peptidase, membrane-bound form #status experimental <MATTM>
F;1-28/Domain: signal sequence #link MATS #status experimental <SIG>
F;1-56/Domain: intracellular #status predicted <INT>
F;7-26/Domain: transmembrane #status predicted <TN>
F;29-72/Domain: extracellular #status predicted <EXT>
F;29-72/Domain: peptidase #link MATS #status experimental <PRO>
F;35-72/Product: dipeptidyl-peptidase, soluble form #status experimental <MATS>
F;83,90,148,217,227,319,421,686/Binding site: carbohydrate (Asn) (covalent) #status pred
F;631/Active site: Ser #status experimental

F;709,741/Active site: Asp, His #status predicted
Query Match 10.9%; Score 510.5; DB 1; Length 792;
Best Local Similarity 25.7%; Pred. No. 4e-28;
Matches 182; Conservative 105; Mismatches 287; Indels 133; Gaps 30;
QY 206 PADPMWIAPIHSND-----IWTSNVTREERRLTYVNEELANMEEDARSAGVATFVLQEB- 260
DB 147 PNTQITITWSEGHKLAVYKNDIYVLEPHLP-SHRTTSTGKENVIFNGINDWVVEEI 205
QY 261 FDRSYGWCPCAETTPSGGKILRLYEENDESEVEIIVHVTSPMLETADSRFRPKCTG 320
DB 206 FCAYSALMW-----SPNGTFLAVAQNDGTGVLPIEYSFYSDLSQYPTVWIPYKACA 259
QY 321 ANPKVTFKMSIMDAEGRIIDVTDKBLIOFFELFEGVEYIARAGWTPEGKYAWSILLD 380
DB 260 VNPVTQKFTI-VNTDSLSTTTTTTTPMOTAPASVT-TGDHYLCDVAWVSEDRIS----- 310
QY 381 RSQTRLQIVLISPFLIFIVEDDVNMRQLIESVPDSVTPLIIYEETTDIWINIHDIFHV- 439
DB 311 -----LQWLRRIQNY--SYMAICDYDKTNLWNCCTTQEHIE 345
QY 440 -----FPQSHHEEIEFIFASECKTGFRHLYKITISILKESKYKRSSGGGLPAPS 486
DB 346 TSATGWGVRFRPAEPHFTSDGSSFYKIVSDKGYKH---ICQFQDKRK----- 390
QY 487 DFKCPTKEBIATTSGEWEVLGRHGSNIQVDEVRRLVYFEGT-KDSPLEHHLHYVSVNPG 545
DB 391 -----PQDCTFITKAGEVI-----SIEALTSDYLYIISNEYKEMPGRNLYKI----- 435
QY 546 EVTRLTDRGYSHSCCIS-----OHCDFP---ISKVSNOKNPHC-----VSLYKLSPPEDDP 593
DB 436 ---QLID--HTNKKLSCDLNPERCOYYSVLSKEAKYVQYOLGCRGCLPLTYTLHRSDQK 490
QY 594 TCKTKFWATILDSAGPIPDYTPPE---LPSFESTTGTLGYMLYKPHDLQPGKYPTVL 650
DB 491 ELRVLED-NSALDKM--LQDVMPKSLDPLVLETRF-WYQMLLPH-PDKSKYPELLI 545
QY 651 FIYGGQVQLVNNRFGVKYFRLN---TLASLGVVVVV-INNRSGHRLKFEAGFYKYM 706
DB 546 DVYAGPCSQ-----KADAAFRFNWATYLASTENIIVASFDRSGYQGGKIMHAINKRL 599
QY 707 GOEIDDDVEGL-QYLASRYDFILDRVGIHGSYGGVLSLMLMQRSDIERVAVIAGPV 765
DB 600 GTLEVEDQIEAARQFL--KMGFVDSKQVAIWGVSYGGVVTSVNLGSGSGVFKCGIAVAPV 657
QY 766 TLMIFYDTGTYERYMG--HPDQNEQGYLGSVAMQAEKFPSEPNRLLHLLHGLDENVHFA 823
DB 658 SRWEYDVSVYTERYMGLTPTEDNLDHYRNSTVMSRAENF--KQVEYLLIHGTADDNVHFQ 715
QY 824 HTSILLSFLVRACKPYDLQIYQERHSIRVPESGEHVELHLLHVLQE 870
DB 716 QSAQISKALVDAGVDFQAMWYTDDEHGIASSTAHOHIYSHMSHFLOQ 762

RESULT 9

A30107
dipeptidyl aminopeptidase B (BC 3.4.14.-) - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein YHR028c
C;Species: Saccharomyces cerevisiae
C;Date: 07-Jun-1990 #sequence_revision 30-May-1997 #text_change 09-Jul-2004
C;Accession: S46780; A30107

submitted to the EMBL Data Library, June 1994

A;Description: The sequence of S. cerevisiae cosmid 8082.

A;Reference number: S46773

A;Accession: S46780

A;Molecule type: DNA

A;Residues: 1-818 <DUZ>

A;Cross-references: UNIPROT:P19962; UNIPARC:UPI0000031ASF; EMBL:U10399; NID:G500689; PI:

R;Roberts, C.J.; Pohlig, G.; Rothman, J.H.; Stevens, T.H.

J. Cell Biol. 108, 1363-1373, 1989

A;Title: Structure, biosynthesis, and localization of dipeptidyl aminopeptidase B, an i

A:Reference number: A30107; MUID:89174971; PMID:2647766
A:Accession: A30107
A:Molecule type: DNA
A:Residues: 1-82, 'H', 84-124, 'N', 126-181, 'LRRLET', 189-199, 'N', 201-365, 'DFKGRKER', 376-574
A:CROSS-references: UNIPARC:UPI0000172A3C; EMBL:X15484
A:Note: the authors translated the codon ACC for residue 572 as Asn
C:Genetics:
A:Gene: SGD:DAP2; STB13; MIPS:YHR028C
A:Map position: 8R
A:Cross-references: SGD:S0001070; MIPS:YHR028C
C:Superfamily: dipeptidyl-peptidase IV
C:Keywords: dipeptidylpeptide hydrolase; glycoprotein; transmembrane protein; yeast vacuole
F:30-45/Domain: transmembrane #status predicted <MM>
F:63,79,110,139,392,421/Binding site: carbohydrate (asn) (covalent) #status predicted

Query Match 10.38; Score 484.5; DB 1; Length 818;
Best Local Similarity 26.08; Pred. No. 3e-26;
Matches 183; Conservative 95; Mismatches 250; Indels 177; Gaps 35;

QY 212 IAFTHSNDIWSNIVTREERLTYVHNE-----LANMEEDARSAGVATFVLQEE-FDRYSG 266
DB 201 IAYQDNIIYISAIKKTIRA--VTNDGSSFLNGKPD-----NVYEEVFEDDKA 250
QY 267 YWCPKAEPTSPGKILRLIYENDSEV-EIHVTSPLMLETTRAD-----SPRYPTG 319
DB 251 AWSPTGD-----YLAFLKIDSEVGEFT---IPYVQDEKDIYPMRSIKYPRSG 298
QY 320 TANPKVFKSEIMIDAGRIIDVIDKELIQP--FEILFEGVEYIARAGWTPEGYAWSI 377
DB 299 TPNP-----HAELWVYSKMDGTGFHPRISNGKKDGLLITETVWVGNG-----NV 343
QY 378 LL-----DRSQRLQLVLSPELFIPEVDVMEORLIESVPDSVTPLLIYEETDIWINI- 433
DB 344 LVKTTDRSSDLITVFLI-----DTAKTSNVVRN-----ESNGGWEIF 383
QY 434 HDIFHV-----FPQSHEEIEFIPASECTGFRHLKYKTSILKESKYRSGGLPAPSDP 489
DB 384 HNTLFI PANETFDPRHNGYVDILPIG---GYNHL-----AYPENSNS-----SHY 425
QY 489 KPCKIEIAITSGEWEVLGRHSGNIQVDEVRRLVPEGTGKDSPLHLYVYVYNPGEV 548
DB 426 K-----TLTEGKWEV---NGPLAFDSMENRLYFISTRKSSSTERHVIYIDLSPNEII 475
QY 549 RLTRDGRYS--HSCCISQHCDFIFISKYSNKNPH-----C-----VSlyKL 586
DB 476 EVTIDSEGVYDVSFSGRFGLLTYGKVPYQKIVDFHSRKAECCKGNVGLKSLYHL 535
QY 587 SSPEDDPTCKTFEATILDSAGPLPDYTPPEIFSPFSTGTFLYG--MLYKPHDLQ-- 642
DB 536 EKNE-----VLTKI--LEDYAVPRKSFRELNLGDKDFGKDLVNSVEILLPND 580
QY 643 -----GKKYPTVLFIYGGPOVLVNNRFGVKYFRNL-----TLASLGYYVWVIDNRSGCH 693
DB 581 FDETLSDHYVPFFFAFGPNSSQV-----VKTFSYGFNEVVASQNALIVWVDVGRGTGF 634
QY 694 RGLPEGAFFKYMGOIIDDQVEGLQVLSRYDFIDLDRVGHGWSYGYLSLMAIMQRS 753
DB 635 KGQDFRSILVRDLGDYIARDQISAAISLYGS-LTFVDPQKISLFGMSYGYLFLKLEKDG 693
QY 754 D-IFRVAIAGAPVTLWIFDYGTERTYMGHPDQNEQGYLYGSV-----AMQAEKFFSEPNR 808
DB 694 GRHFKYGNVAPVDMRFYDYSVTERYMYHTPOENFDGYESSVHNVTL-----AQANR 747
QY 809 LLLHGLFDENVHFAHTSILLSLFLVRAG-KPYDLQIYPOERHSIR 852
DB 748 FLAMHGTGDDNVHFQNSKFLDLIDLNGVENYDVHVPDSDHSIR 792

RESULT 10
A49737
dipeptidyl aminopeptidase (EC 3.4.14.-) - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein O5045; protein YOR219c; protein YOR50-9
C:Species: Saccharomyces cerevisiae

C:Date: 03-Mar-1994 #sequence_revision 03-Mar-1994 #text_change 09-Jul-2004
A:Accession: A49737; S45451; S60946; S67112; S71721
R:Santa Anna-A, S.; Herskowitz, I.
submitted to the Protein Sequence Database, July 1993
A:Reference number: A49737
A:Accession: A49737
A:Molecule type: DNA
A:Residues: 1-931 <SAN>
A:Cross-references: UNIPROT:P33894; UNIPARC:UPI0000136060; GB:L21944; NID:g347196; PIDN:PIDN
R:Anna-Ariola, S.S.; Herskowitz, I.
Yeast 10, 801-810, 1994
A:Title: Isolation and DNA sequence of the STE13 gene encoding dipeptidyl aminopeptidase
A:Reference number: S45451; MUID:95066382; PMID:7975897
A:Accession: S45451
A:Molecule type: DNA
A:Residues: 1-931 <ANN>
A:Cross-references: UNIPARC:UPI0000136060; EMBL:L21944; NID:g347196; PIDN:AAA35119.1; P
R:Gallison, F.; Dujon, B.
submitted to the EMBL Data Library, October 1995
A:Description: Sequence and analysis of a 33 kb fragment from the right arm of chromosome
A:Reference number: S60938
A:Accession: S60946
A:Molecule type: DNA
A:Residues: 1-931 <GL>
A:Cross-references: UNIPARC:UPI0000136060; EMBL:X92441; NID:g1050762; PIDN:CAA63182.1;
R:Boyer, J.; Fairhead, C.; Gallison, L.; Gallison, F.; Michaux, G.; Thierry, A.; Dujon,
submitted to the Protein Sequence Database, July 1996
A:Reference number: S67104
A:Accession: S67112
A:Molecule type: DNA
A:Residues: 1-931 <BO>
A:Cross-references: UNIPARC:UPI0000136060; EMBL:Z75127; NID:g1420507; PIDN:CAA99437.1;
A:Experimental source: strain S288C
R:Gallison, F.; Dujon, B.
Yeast 12, 877-885, 1996
A:Title: Sequence and analysis of a 33 kb fragment from the right arm of chromosome XV
A:Reference number: S71713; MUID:96437977; PMID:8840505
A:Accession: S71721
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-931 <GAW>
A:Cross-references: UNIPARC:UPI0000136060; EMBL:X92441; NID:g1050762; PIDN:CAA63182.1;
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1995
C:Genetics:
A:Gene: SGD:STE13; YCI1
A:Cross-references: SGD:S0005745; MIPS:YOR219c
A:Map position: 15R
C:Function:
A:Description: involved in processing of alpha-factor prepropheromone
C:Keywords: dipeptidylpeptide hydrolase; glycoprotein; transmembrane protein
F:113-150/Domain: transmembrane #status predicted <TM>
F:377/Binding site: carbohydrate (asn) (covalent) #status predicted
F:785,863,896/Active site: Ser, Asp, His #status predicted

Query Match 9.94; Score 465; DB 2; Length 931;
Best Local Similarity 21.54; Pred. No. 9e-25;
Matches 218; Conservative 150; Mismatches 380; Indels 268; Gaps 41;

QY 6 ETEQLGVEIFTADCEENIESQBPKEPPEYVRYSWSQLKLLADTRKYHGYMAKAPH 65
DB 43 OSNENGHTINIRTEATIDVTDPQ--TPFLQOYS-----MRPRE 83
QY 66 DMFVKRNDPGHSDRIYVYLAWS-----GENRENTLFYSEIPKTNIRAAVLMIS 115
DB 84 SFQF---NDIENQHHTHSFFSVNKFNRWGEWSLPKRSVLVFTLLIALSVLLVILIP 140
QY 116 WKPL-----LDLFOATL-----DYGM 131
DB 141 SKLLPTKITRPTKSAGDSSLGRSFSIENVLNGDFAIPEDTFHFIDPQRLGLQSDPGL 200
QY 132 ISREELLREKRIG-TVGASDYDHYHGSCTFLFQ----- 165
DB 201 YPTTEIDGHTNFTIAKQLFDETFEVNIGGNRFLYEGVEFTVTVQINLYKLDKILFTNL 260

C;Superfamily: dipeptidyl-peptidase IV		Query Match	9.6%; Score 451.5; DB 2; Length 799;
Best Local Similarity		23.0%;	Fred. No. 6.5e-24;
Matches 195; Conservative 129; Mismatches 298; Indels 227; Gaps 38;			
Qy	78	PHSDRIYILAMS-----GENRENTLFYSEIPKTIINRAAVLMLSWKPLL-----DLF	123
Db	116	PSADRKYFAMDHAFNPGMNFQNETFHLKIVNNERL-----NPLPFEVEELPRELS	168
Qy	124	QATLDYGMYSREELLREKRIGTVGIASVDYHOGSGTFLFOAGSGIYHVKGDPGQFTQ	183
Db	169	DSRITTYDILGRKEESV-----IOAFKWKNGKFNDFVFNESNKIYY-----	207
Qy	184	QPLRPNLVETSCPNIRMDPKLCPADPDWIAFIHSNDIWSINIVTREERRITYVHNELANM	243
Db	208	-----QSSP-----ELEGTRVNSNGEHT	226
Qy	244	EEDARSAGVATFVLOEE-FDRYSGYWCPCKAETTPSGGKILRLILYEENDESEVEIIVHTS	302
Db	227	VD-----GLFDWIYEEEIFGRKDAWMSWK-----GDQLAYASYDNHLTKVSL--KTY	273
Qy	303	PMLSTRAD-SFRYPKT-----GTANPKVTFKMSIEMIDABGRIIDVIDIKELIOPPEIL-	355
Db	274	HLLEPYDITNTHYFKTAKVLPYTLISLWKKTE-----QSRQLDVQLKDSLSYHYLLA	328
Qy	356	-----PEGVEYIARAGWTPEGKYAWSILL-----DRSOTRLOIIVLISPELLFIPVEDDVM	405
Db	329	VKWEINGTEQLVSV-WT--NRIQNEVALTICDWDTAICRLEPE-----	369
Qy	406	RORLIESVDSVTPLLIYEETDDIWINIHDIHFVFPQSHHEEIEFIFASECKTGFRHLVK	465
Db	370	-----YKASKRWVT-HDDFHSI-TSFEDTLFLLP-----HDKR	402
Qy	466	ITSILKESKYKSSGGLPAPSPDFKCPICEIAITSGEWELGRHGSNIQWDEVRLVYPE	525
Db	403	DNAFOQVASLRLSHGQLRTPK-----FLNLGEYDVTSINGINKET-----RTIFFH	448
Qy	526	GTKDSPLEHLYVVSVPNGEVTRLTDRGYSHSCIS---QHCDFFISKYSNQNPHCVS	582
Db	449	AAAPKPSHRSLSFS-----LADESRNSAYCISCIKNCTMAQADDDQMKTAIVS	499
Qy	583	LYKLSGSPED-----DPTCKTK-----EFWATILDSAGPLDPYTPPEIFSFTTG	627
Db	500	CKGPAAPHTATVNLTRMDSDKKTEHANLLYDKTYQNRVEEAG-LFVIAIKETI---KISDD	555
Qy	628	FTLYGMLYKPHDLQPGKKY---PTVLFIYGGPQVQVNNRPFKGVKFRNLNTLASLGYVV	684
Db	556	FDALIKLSIPKDIYNRDKHQAIPLIHVYVGGNDQ---NTKEATQIGIEEVVASASQAAI	612
Qy	685	V-IDNRGSGCHRGKFEKGFAGKYKMGQIEIDQVQEGLOYLASRY-DFIDLDVRGIHGWSYGG	742
Db	613	LRIDGRSGGRGWKYSALYQGLGIVEVEDQIKAIKVVLLRYRHLLDARRVAVFGWSYGG	672
Qy	743	YLSLMALMQRSD-IFRVAIAGAPVTLMIFYDTGYTRYMGHPDQNEQGYILGVSAMQAEK	801
Db	673	FMTLSMNEAPEOFFKCAVSAPVTFNFAVDATYTYRWMG--DAPLESY--SDVTKKLDN	728
Qy	802	FPSPENRLLHLHGFLENVHFAHTSILLSFVRAGKPDLOIYPOERHSIRVPSSGEHYE	861
Db	729	FKS--TELLMHGLLDNHNHVFQNSALLDELQNRGVDFDLMVIFNQAHSLSRYS--HVV	784
Qy	862	LHLHLYLOE 870	
Db	785	GKMTFLRQ 793	
RESULT 12			
T25173			
hypoetical protein T23F1.7a - Caenorhabditis elegans			
C;Species: Caenorhabditis elegans			
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004			
C;Accession: T25173			

Qy	166	-----AGSGIYHVKGDPGQFTQOPLRPNLVETSCPNIRMDPK-----LCRAD	208
Db	261	SEFRHSSKGYWIKD-----LNTGNIEPILPPEKSDDDNVELGLSKLSYAHFSPA-	309
Qy	209	PDWIAFIHNDIWSINIVTREERRITY-VHNELANMEEDARSAGVATFVLOEE-FDRYSG	266
Db	310	YNYIYFVYENNLFLQOVNSGVAKKVTEDGSKDIENAKPD-----WIYEEVLASDOA	361
Qy	267	YWCPCKAETTPSGGKILRLILYEENDESEVEIIVHTSPMLSTRADS-FRYPKTGTANPKV	325
Db	362	IWW-----APDSCAVAFARFNDISVDIIRLNRYTN-MNEAYLSDTKIKYKPGFQNPQF	414
Qy	326	TFKMSIEMIDABGRIIDVIDIKELIOPPEILPEGVEYIARAGWTPEGKYAWSILLDRSQR	385
Db	415	DL-----FLVNLQNGIYISINTG-GOKDSILYNG-----KWISPDTFPEI-TDRNSKI	461
Qy	386	LOIVLSPELFIPVEDDVMERQRLIESVDSVTPLLIYEETDDIWINIHDIHFVFPQSH	444
Db	462	LDVKVY-----DIPSSQML--TVRNTSNLF-----NGWIEKTKDILISIPKPE	503
Qy	445	EEBIE--FIFASECKTGFRHLKYITSILKESKYKSSGGLPAPSPDFKCPICEIAITSGE	502
Db	504	LKRMDCYGIIDIHADSRGFSHLFYPTVF-----AKEPIQLTKGN	542
Qy	503	WEVLGRH--GSIQIWDVEARLVYFEGTKDPSLEHLYVVSV-----NPGEV	548
Db	543	WEVTNGIVGYEYETD---TIFFTANEIGVMSQRLYSISLTDSTTQNTQTSQONP----	594
Qy	549	RLTRDGYSHSCISQHCDFIFISKYSNQNKP-----HCVSLYKLSGSPEDD	592
Db	595	--SDKYDFYDFELSSSARVAISKLGPDTEIKVAGPLTRVLNVAEIHDDSILOTKDE--	650
Qy	593	PTCKTEFWATILDSAGPLDPYTPPEIFSFTT---GFTLYGMLYKPHDLQPGKKYPTV	649
Db	651	--KFXE-----KIKNYDLP-ITSYKTMVLDDGVEINYIEIKPANLNPKKYPIL	696
Qy	650	LFIYGGPQVQVNNRPFKGVKFRNLNTLASLGVVVVVDNRGSGCHRGKFEKGFAGKYKMGQI	709
Db	697	VNIYGGSGQFTT--KSLAFEQAVSGLDVIVLQIEPRGTGKGKGSFRSWAREKLGWY	754
Qy	710	EIDDOVE-GLQYLASRYDFIDLDVRGIHGWSYGGVLSMAL--MORSDFRVAIAGAPVTL	767
Db	755	EPKIDITEVTKKTFORNSQHIDSKAIWGSYGGFTSLKTYVELDNGDTFKYMAVAVTN	814
Qy	768	WIFYDTGYTRYMGHPDQNEQGYILGVSAMQAEKTPSPENRLLHLHGFLENVHFAHTSI	827
Db	815	WTLYDSVYTRYVWNPQSENHEGYFEVSTIQNFKFS-LEKLFIVHGTFFDNVHIQNTFR	873
Qy	828	LLSFLVRAG-KPYDLOIYPOERHSIRVPSSGEHYELHLLHYLQENLGSRIAALKVI	882
Db	874	LVDQNLGLTNYDMHIFPDSDHSIRYHNAQRIVFQKLYYWLDAFERFDNTEVL	929
RESULT 11			
T25174			
hypoetical protein T23F1.7b - Caenorhabditis elegans			
C;Species: Caenorhabditis elegans			
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004			
C;Accession: T25174			
R;Wilkinson, J			
submitted to the EMBL Data Library, October 1996			
A;Reference number: 219990			
A;Accession: T25174			
A;Status: preliminary; translated from GB/EMBL/DBJ			
A;Molecule type: DNA			
A;Residues: 1-799 <WIL>			
A;Cross-references: UNIPROT:O18119; UNIPARC:UPI0000061190; EMBL:Z81129; PIDN: CAB03412.1			
A;Experimental source: clone T23F1			
C;Genetics:			
A;Gene: CESP:T23F1.7b			
A;Map position: 5			
A;Introns: 13/3; 52/1; 111/3; 197/3; 366/3; 524/1; 557/1; 648/2; 703/1; 743/1			

Db 723 LVITHATDEKIHFOHTAELITQLIRKANKYSLQIYPDESHYFSAALQOHLHRSILGPFV 782
QY 870 E 870
Db 783 E 783
RESULT 14
I68600
dipeptidyl aminopeptidase like protein - human
C:Species: Homo sapiens (man)
C>Date: 06-Sep-1996 #sequence_revision 06-Sep-1996 #text_change 21-Jul-2000
C/Accession: I68600
R:Yokocani, N.; Doi, K.; Wenthold, R.J.; Wada, K.
Hum. Mol. Genet. 2, 1037-1039, 1993
A:Title: Non-conservation of a catalytic residue in a dipeptidyl aminopeptidase IV-relat
A:Reference number: I54331; MUID:93372805; PMID:8103397
A:Accession: I68600
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-803 <RES>
A:Cross-references: UNIPARC:UPI000016A7F9; GB:M96860; NID:G306707; PIDN:AAA35761.1; PID:
C:Superfamily: dipeptidyl-peptidase IV
Query Match 9.1%; Score 427; DB 2; Length 803;
Best Local Similarity 23.4%; Pred. No. 3.7e-22;
Matches 172; Conservative 121; Mismatches 299; Indels 144; Gaps 30;
QY 171 YHVKGDPGPGFTQQPLRPNLVETSCPNIRMDPKLCPADPDWIAFIHSNDIWSINVTREE 230
Db 156 YVLSKIPHGDPQSLDPEVSNAKLQYAGWGPK-----GQQLIFIPENNIYYCAHVKGQA 210
QY 231 RRLTYVHNLANNEEDARSAGVATFVLOEE--FDYSGYVWCPCKAETTPSGGKILRLIYEE 289
Db 211 IRV-----VSTGKEGVIYNGLSDWLYEEELKTHIAHWSPDG-----TRLAYAA 255
QY 290 NDESEVEITHV---TSPMLETRRADSFYPKTGTANPKVTFKMESEIMDAEGRIIDVIDK 346
Db 256 INDSRVPIMLPTTYSIYPT--VKPYHYPKAGSENPSISLH-----VIGLNGPTHDL--- 306
QY 347 ELIQPEILLFEGVEYIARAGWTPEGKYA--WSILLDRSQRLQIQLVLSPELFIPEVDVDM 404
Db 307 EMPPDDPRMREY-YITMVKWATSTKVAVTW---LNRAQN-----VS----- 344
QY 405 ERORLIESVDSVTPLII--YEETDDIWINIHDIHFVFPQSHREEI-----EFIPASEC 456
Db 345 -----ILTLCDAFTGVCTKKHEDESEAMLH-----RQNEBPVFSKDGKFFIRAI 390
QY 457 KTGPR-HLYKITSLKESKYKSSGGLPAPSPDKPIKEEIAITSGGEWVLGRHGSNIQV 515
Db 391 PQGGRGKFYHIT--VSSOPNSNDNIQ-----SITSGDMDVT---KILAY 431
QY 516 DEVRLVYFGTKDPSLEHLLYVSVYVNPGEVTRLTDRGYSH---SCCISOHCDFIFISKY 572
Db 432 DEGNKIYFLSTEDLPRRRLQYSAN-----TEGNFNRQCLSCDLVENCITYFSASF 481
QY 573 SNQKNPHCVSLYKLSGP-----EDDPTCKTEF-----WATILDSAGPLDPDTPPE 618
Db 482 SHSMD-----FELLKCEGPGVPMVTVHTTDDKKQMFLETNEHVKKAINDRQMPKVEYRDIE 538
QY 619 IFSPESTGTFTLYGMYLKPHDLQPGKKYPTVLIYGGPQVQLVNNRFGVKYFRLNLTAS 678
Db 539 I-----DDYNLPMQLIKPATFTDTHYPLLLVVDGTGSGSVAEKFE--VSWETVMVSS 590
QY 679 LGYVVVVINDRGSCHRGKPEGAFKYMGOIIEIDQVGLQVLAARYDFIDLDVRGIGHW 738
Db 591 HGAVVVKCDGSGSGFQGTLLHVEVRRRLGLEEKQMEAVRML--KEQYIDRTVAVFGK 649
QY 739 SYGGLYSIMALMQRSD-----IFRVAITAGAPVTLWIFDYDTGYTERYMGHPDQNEQGYLGS 794
Db 650 DYGGYLTSTYLPAKGENQGTFTCGSALSPITDFDKLYASAFSEYVILGLHGLDNRAYEMTK 709
QY 795 VMAQAEKPEPSEPNRLLLHGLFDENVHFPAHTSILLISFLVRACKPYDLQIYQERHSIRVP 854

Db 710 VAHRVSAL--BEQQFLIHIPTADEKIHFOHTAELITQLIRKANKYSLQIYPDESHYFTSS 767
QY 855 ESGEHVELHLHYLOE 870
Db 768 SLKOHLYRSIINFFVE 783
RESULT 15
I54331
dipeptidyl aminopeptidase like protein - human
C:Species: Homo sapiens (man)
C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C/Accession: I54331
R:Yokocani, N.; Doi, K.; Wenthold, R.J.; Wada, K.
Hum. Mol. Genet. 2, 1037-1039, 1993
A:Title: Non-conservation of a catalytic residue in a dipeptidyl aminopeptidase IV-rela
A:Reference number: I54331; MUID:93372805; PMID:8103397
A:Accession: I54331
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-865 <RES>
A:Cross-references: UNIPROT:P42658; UNIPARC:UPI0000047378; GB:M96859; NID:G306705; PIDN:
C:Superfamily: dipeptidyl-peptidase IV
Query Match 9.1%; Score 427; DB 2; Length 865;
Best Local Similarity 23.4%; Pred. No. 4.1e-22;
Matches 172; Conservative 121; Mismatches 299; Indels 144; Gaps 30;
QY 171 YHVKGDPGPGFTQQPLRPNLVETSCPNIRMDPKLCPADPDWIAFIHSNDIWSINVTREE 230
Db 218 YVLSKIPHGDPQSLDPEVSNAKLQYAGWGPK-----GQQLIFIPENNIYYCAHVKGQA 272
QY 231 RRLTYVHNLANNEEDARSAGVATFVLOEE--FDYSGYVWCPCKAETTPSGGKILRLIYEE 289
Db 273 IRV-----VSTGKEGVIYNGLSDWLYEEELKTHIAHWSPDG-----TRLAYAA 317
QY 290 NDESEVEITHV---TSPMLETRRADSFYPKTGTANPKVTFKMESEIMDAEGRIIDVIDK 346
Db 318 INDSRVPIMLPTTYSIYPT--VKPYHYPKAGSENPSISLH-----VIGLNGPTHDL--- 368
QY 347 ELIQPEILLFEGVEYIARAGWTPEGKYA--WSILLDRSQRLQIQLVLSPELFIPEVDVDM 404
Db 369 EMPPDDPRMREY-YITMVKWATSTKVAVTW---LNRAQN-----VS----- 406
QY 405 ERORLIESVDSVTPLII--YEETDDIWINIHDIHFVFPQSHREEI-----EFIPASEC 456
Db 407 -----ILTLCDAFTGVCTKKHEDESEAMLH-----RQNEBPVFSKDGKFFIRAI 452
QY 457 KTGPR-HLYKITSLKESKYKSSGGLPAPSPDKPIKEEIAITSGGEWVLGRHGSNIQV 515
Db 453 PQGGRGKFYHIT--VSSOPNSNDNIQ-----SITSGDMDVT---KILAY 493
QY 516 DEVRLVYFGTKDPSLEHLLYVSVYVNPGEVTRLTDRGYSH---SCCISOHCDFIFISKY 572
Db 494 DEGNKIYFLSTEDLPRRRLQYSAN-----TEGNFNRQCLSCDLVENCITYFSASF 543
QY 573 SNQKNPHCVSLYKLSGP-----EDDPTCKTEF-----WATILDSAGPLDPDTPPE 618
Db 544 SHSMD-----FELLKCEGPGVPMVTVHTTDDKKQMFLETNEHVKKAINDRQMPKVEYRDIE 600
QY 619 IFSPESTGTFTLYGMYLKPHDLQPGKKYPTVLIYGGPQVQLVNNRFGVKYFRLNLTAS 678
Db 601 I-----DDYNLPMQLIKPATFTDTHYPLLLVVDGTGSGSVAEKFE--VSWETVMVSS 652
QY 679 LGYVVVVINDRGSCHRGKPEGAFKYMGOIIEIDQVGLQVLAARYDFIDLDVRGIGHW 738
Db 653 HGAVVVKCDGSGSGFQGTLLHVEVRRRLGLEEKQMEAVRML--KEQYIDRTVAVFGK 711
QY 739 SYGGLYSIMALMQRSD-----IFRVAITAGAPVTLWIFDYDTGYTERYMGHPDQNEQGYLGS 794
Db 712 DYGGYLTSTYLPAKGENQGTFTCGSALSPITDFDKLYASAFSEYVILGLHGLDNRAYEMTK 771

Qy 795 VAMQAEKFPSEPNRLLLHGFLENNVHFAHTSILLFLVRAGKPYDLOIYPQERHSIRVP 854
Db 772 VAHRVSAL--EEQFLIIHPTADEKIHFOHTAELITQLIRGKANYSLQIYPDESHYFTSS 829
Qy 855 ESCEHYELHLLHYLQE 870
Db 830 SLKQHLYSIINFVE. 845

Search completed: April 14, 2006, 12:51:03
Job time : 55 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 14, 2006, 12:43:37 ; Search time 240 Seconds
(without alignments)
2592.817 Million cell updates/sec

Title: US-10-825-632-1
Perfect score: 4700
Sequence: 1 MAAMETQLGVETADTC.....HLLHYLQENLGRRAALKVI 882

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 2166443 seqs, 705528306 residues
Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Databases: Uniprot_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	4700	100.0	898	1	DPB8_HUMAN	Q6V1X1 homo sapien
2	4336.5	96.5	892	1	DPB8_MOUSE	Q80YV7 mus musculus
3	2870	61.1	863	1	DPB9_HUMAN	Q86T12 homo sapien
4	2833	60.3	862	1	DPB9_MOUSE	Q8BV94 mus musculus
5	2829.5	60.2	923	2	Q4SBM6_TETNG	Q4sbm6 tetraodon n
6	2707.5	57.6	847	2	Q6GR22_XENLA	Q6gr22 xenopus lae
7	1643.5	35.0	508	2	Q75273_HUMAN	Q75273 homo sapien
8	1641	34.9	886	2	Q70BK1_ANOGA	Q70bk1 anopheles g
9	1596	34.0	1053	2	Q9VC20_DROME	Q9vc20 drosophila
10	1596	34.0	1113	2	Q9VC19_DROME	Q9vc19 drosophila
11	1567	33.3	740	2	Q5TTK8_ANOGA	Q5ttk8 anopheles g
12	1321	28.1	621	2	Q7PTT8_ANOGA	Q7ptt8 anopheles g
13	1142.5	24.3	557	2	Q5TXJ2_ANOGA	Q5txj2 anopheles g
14	1124.5	23.9	803	2	Q54U01_DICTDI	Q54u01 dictyosteli
15	1049	22.3	432	2	Q75868_HUMAN	Q75868 homo sapien
16	924.5	19.7	931	2	Q965K3_CAEEL	Q965k3 caenorhabdi
17	922.5	19.6	927	2	Q44987_CAEEL	Q44987 caenorhabdi
18	867.5	18.5	895	2	Q61CU7_CAEER	Q61cu7 shewanella
19	846	18.0	763	2	Q8EAB7_SHEON	Q8eab7 shewanella
20	842	17.9	738	2	Q9A6E0_CAUCR	Q9a6e0 caulobacter
21	837.5	17.8	746	2	Q9FNF6_ARATH	Q9fnf6 arabidopsis
22	792.5	16.9	745	2	Q6F317_9PSED	Q6f317 pseudomonas
23	786	16.7	596	2	Q6K880_ORISA	Q6k880 oryza sativ
24	779	16.6	743	2	Q5QX36_IDILO	Q5qx36 idiomarina
25	759	16.1	745	2	Q5H5W8_XANON	Q5h5w8 xanthomonas
26	754	16.0	741	2	P95782_XANNA	P95782 xanthomonas
27	751.5	16.0	751	2	Q4UPD3_XANCP	Q4upd3 xanthomonas
28	751.5	16.0	751	2	Q8P3V8_XANCP	Q8p3v8 xanthomonas
29	746	15.9	766	2	Q4TNP1_9SPHN	Q4tnp1 erythrobact
30	744.5	15.8	757	2	Q8PFD7_XANAC	Q8pfd7 xanthomonas
31	727	15.5	552	2	Q8GUJ7_ARATH	Q8guj7 arabidopsis

RESULT 1				
ID	DPB8_HUMAN	STANDARD;	PRT;	898 AA.
AC	Q6V1X1; Q7Z4C8; Q7Z4D3; Q7Z4E1; Q81WG7; Q8NEM5; Q96JX1; Q9HEM2;			
AC	Q9HEM3; Q9HEM4; Q9HEM5; Q9NXP4;			
DT	13-SEP-2005 (Rel. 48, Created)			
DT	13-SEP-2005 (Rel. 48, Last sequence update)			
DT	13-SEP-2005 (Rel. 48, Last annotation update)			
DE	Dipeptidyl peptidase 8 (EC 3.4.14.5) (Dipeptidyl peptidase VIII) (DP8)			
DE	(PolyI dipeptidase DPP8) (Dipeptidyl peptidase IV-related protein 1) (DPP8-1).			
GN	Name=DPP8; Synonyms=MSTP097, MSTP135, MSTP141;			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;			
OC	Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 3), NUCLEOTIDE SEQUENCE [MRNA] OF			
RP	334-898 (ISOFORM 4), NUCLEOTIDE SEQUENCE [MRNA] OF 540-898 (ISOFORM			
RP	5), NUCLEOTIDE SEQUENCE [MRNA] OF 260-792 (ISOFORM 6), FUNCTION,			
RP	CATALYTIC ACTIVITY, ENZYME REGULATION, TISSUE SPECIFICITY, INDUCTION,			
RP	AND SUBCELLULAR LOCATION.			
RC	TISSUE=Placenta;			
RX	MEDLINE=20467194; PubMed=11012666;			
RA	Abbott C.A., Yu D.M.T., Woollatt E., Sutherland G.R., McCaughan G.W.,			
RA	Gorrell M.D.;			
RT	"Cloning, expression and chromosomal localization of a novel human			
RT	dipeptidyl peptidase (DPP) IV homolog, DPP8.";			
RL	Eur. J. Biochem. 267:6140-6150(2000).			
RN	[2]			
RP	NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 3), CATALYTIC ACTIVITY, ENZYME			
RP	REGULATION, BIOPHYSICOCHEMICAL PROPERTIES, SUBCELLULAR LOCATION, AND			
RP	TISSUE SPECIFICITY.			
RC	TISSUE=Testis;			
RX	PubMed=12662155; DOI=10.1042/BJ200201914;			
RA	Qi S.Y., Riviere P.J., Trojnar J., Junien J.-L., Akinsanya K.O.;			
RT	"Cloning and characterization of dipeptidyl peptidase 10, a new member			
RT	of an emerging subgroup of serine proteases.";			
RL	Biochem. J. 373:179-189(2003).			
RN	[3]			
RP	NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1).			
RC	TISSUE=Testis;			
RA	Sha J.H., Zhou Z.M., Li J.M.;			
RL	Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.			
RN	[4]			
RP	NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 4), AND NUCLEOTIDE			
RP	SEQUENCE [LARGE SCALE MRNA] OF 211-898 (ISOFORM 2).			
RC	TISSUE=Hepatosoma, and Placenta;			
RX	PubMed=14702039; DOI=10.1038/ng1285;			
RA	Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,			
RA	Wakamatsu A., Hayaashi K., Sato H., Nagai K., Kimura K., Makita H.,			
RA	Sekine M., Ohayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,			
RA	Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,			

Q5nmh8 zymomonas m
Q7nek8 gloebacter
Q8a028 bacteroides
Q3fyg6 gluconobact
Q6h9e3 trypanosoma
Q5lgu5 bacteroides
Q64xp9 bacteroides
Q4qlh9 leishmania
Q47900 flavobacter
Q51zp7 magnaporthe
Q5qvyr idiomarina
Q4nvvi anaeromyxob
Q42812 aspergillus
Q61872 prevotella

RL Mol. Cell. Proteomics 3:279-286(2004).
CC - FUNCTION: Dipeptidyl peptidase that cleaves off N-terminal
CC dipeptides from proteins having a Pro or Ala residue at position
CC 2. May play a role in T-cell activation and immune function (By
CC similarity).
CC - CATALYTIC ACTIVITY: Release of an N-terminal dipeptide, Xaa-Yaa-|-
CC zaa- from a polypeptide, preferentially when Yaa is Pro, provided
CC zaa is not a hydroxyproline.
CC - ENZYME REGULATION: Inhibited by zinc. Inhibited by the serine
CC proteinase inhibitor 4-(2-aminocetyl)benzenesulphonyl fluoride
CC (ABSF), and by di-isopropyl fluorophosphate. Specifically inhibited
CC by isobutylene derivatives (By similarity).
CC - SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC - SIMILARITY: Belongs to the peptidase S9B family. DPPIV subfamily.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; AK016546; BAB30295.2; -; mRNA.
CC EMBL; BC043124; AAH43124.1; -; mRNA.
CC EMBL; BC059222; AAH59222.1; -; mRNA.
CC MROPS; S09.018; -.
CC Ensembl; ENSMUSG0000032393; Mus musculus.
CC MGI; MGI:1921638; Dpp8.
CC InterPro; IPR001375; Peptidase S9.
CC InterPro; IPR002469; Peptidase S9B.
CC InterPro; IPR000379; Ser esters.
CC Pfam; PF00930; DPPIV_N; 1.
CC Pfam; PF00326; Peptidase S9; 1.
CC Amino acid peptidase; Hydrolyase; Phosphorylation; Protease; Serine protease.
CC ACT_SITE 749 749 Charge relay system (By similarity).
CC ACT_SITE 827 827 Charge relay system (By similarity).
CC ACT_SITE 859 859 Charge relay system (By similarity).
CC ACT_SITE 325 325 Phosphotyrosine.
CC MOD_RES 87 87 G - R (in Ref. 1).
CC CONFLICT 87 87 G - R (in Ref. 1).
CC SEQUENCE 892 AA; 102186 MW; 59081CD9792E03ED CRC64;
Query Match 96.5%; Score 4536.5; DB 1; Length 892;
Best Local Similarity 95.8%; Pred. No. 0;
Matches 846; Conservative 18; Mismatches 18; Indels 1; Gaps 1;
1 MAAMATEQLGVEIFETADCEP-NIESQDRKLEPFYVERYSWOLKLLADTRKYGYM 59
10 MAAMATEQLGVEIFETAECEGNGESQDRPALPEFFYVERYSWOLKLLADTRKYGYM 69
60 MAKAPHDFMFKRNDPDGPHSDRIYYLWAMSGENRENTLIFYSEIKTINRAAVIMLWKPL 119
70 MAKAPHDFMFKRNDPDGPHSDRIYYLWAMSGENRENTLIFYSEIKTINRAAVIMLWKPL 129
120 LDLFQATLDYGMYSREELLRERKRGITVGTASVDYHOGSGTFLFQAGSGIYHVKGGPQ 179
130 LDLFQATLDYGMYSREELLRERKRGITVGTAAVDYHFGSGTFLFQAGSGIYHVKGGPH 189
180 GFTQOPLRNLVTSFCNIRMDPKLCPADPDWIAFHSNDIWIISNIVTREERLTYVHNE 239
190 GFTQOPLRNLVTSFCNIRMDPKLCPADPDWIAFHSNDIWIISNIVTREERLTYVHNE 249
240 LANNEEDARSAGVATFVLOEEDRYSGYWMWCPKATTPSGKILRLIYLENDESEVEI 299
250 LANNEEDARSAGVATFVLOEEDRYSGYWMWCPKATTPSGKILRLIYLENDESEVEI 309
300 VTSMLTETRADSFYKPTGTANPKVTFKMSIMIDAGRIIDVIDKELIOPFELPEGV 359
310 VTSMLTETRADSFYKPTGTANPKVTFKMSIIVDAAGIIVDIDKELVQPFELPEGV 369
360 EYIARAGWTGKYSANILLDRSQTLQIVLISPELFTPVDDVWERORLIESVDSVTP 419
370 EYIARAGWTGKYSANILLDRSQTLQIVLISPELFTPVDDVWERORLIESVDSVTP 429
420 LIIYEETTDIWINIHDFHVPQSHHEEIEFIFASECKTGFPHLYKITSILKESYKRSS 479

Db 430 LIIYEETTDIWINIHDFHVPQSHHEEIEFIFASECKTGFPHLYKITSILKESYKRSS 489
Qy 480 GGLPAPSDFKPIKEETAITSGEVLGRHSGNIQVDEVRVLYYFEGTKDSPLEHLVYV 539
Db 490 GGLPAPSDFKPIKEETAITSGEVLGRHSGNIQVDEVRVLYYFEGTKDSPLEHLVYV 549
Qy 540 SYVNPGEVTRLDTRGYSHSCISCHOCFFISKYSNQNPHCVSLYKLSPEDDPTCKTKE 599
Db 550 SYVNPGEVTRLDTRGYSHSCISCHOCFFISKYSNQNPHCVSLYKLSPEDDPTCKTKE 609
Qy 600 FWATILDSAGLPDYTPPEIPSPFESTTGTFLYOMLYKPHDLQPKKYPTVLVYGGPQV 659
Db 610 FWATILDSAGLPDYTPPEIPSPFESTTGTFLYOMLYKPHDLQPKKYPTVLVYGGPQV 669
Qy 660 LYNNRFKGVKFRNLTLASGLYVVVVDNRGSHRGLKPEGAPKYKMGQIEIDDDVEGLQ 719
Db 670 LYNNRFKGVKFRNLTLASGLYVVVVDNRGSHRGLKPEGAPKYKMGQIEIDDDVEGLQ 729
Qy 720 YLASRYDFIDLDVRGIGHGWSYGGYLSLMALMQRSDFRVAIAGAPVTLWIFYDYGTERY 779
Db 730 YLASRYDFIDLDVRGIGHGWSYGGYLSLMALMQRSDFRVAIAGAPVTLWIFYDYGTERY 789
Qy 780 MGHPPDNEQGYLGSVAMQAEKFPSEPNRLLHLLHVLQENLGSRIAALKVI 882
Db 790 MGHPPDNEQGYLGSVAMQAEKFPSEPNRLLHLLHVLQENLGSRIAALKVI 892
Qy 840 DIQIYPOERHSTRVPESGEHVELHLLHVLQENLGSRIAALKVI 882
Db 850 DIQIYPOERHSTRVPESGEHVELHLLHVLQENLGSRIAALKVI 892
RESULT 3
DPP9 HUMAN STANDARD; PRT; 863 AA.
ID DPP9 HUMAN Q66T12; Q6A137; Q6ALU0; Q6ZMT2; Q6ZNU5; Q8N2J7; Q8N3F5; Q8WXD8;
AC Q66T12; Q6A137; Q6ALU0; Q6ZMT2; Q6ZNU5; Q8N2J7; Q8N3F5; Q8WXD8;
AC Q66T12; Q6A137; Q6ALU0; Q6ZMT2; Q6ZNU5; Q8N2J7; Q8N3F5; Q8WXD8;
DT 13-SEP-2005 (Rel. 48, Created)
DT 13-SEP-2005 (Rel. 48, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Dipeptidyl peptidase 9 (EC 3.4.14.5) (Dipeptidyl peptidase IX) (DPP9)
DE (Dipeptidyl peptidase-like protein 9) (DPLP9) (Dipeptidyl peptidase
DE IV-related protein 2) (DPRP-2).
GN Name=DPP9; Synonyms=DPRP2;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1), AND TISSUE SPECIFICITY.
RX MEDLINE=22347328; PubMed=12459266; DOI=10.1016/S0378-1119(02)01059-4;
RA Olsen C., Wagtmann N.;
RT "Identification and characterization of human DPP9, a novel homologue
RT of dipeptidyl peptidase IV.";
RL Gene 299:185-193(2002).
[2]
RN NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1), CATALYTIC ACTIVITY,
RN BIOPHYSICO-CHEMICAL PROPERTIES, ENZYME REGULATION, TISSUE SPECIFICITY,
RN AND SUBCELLULAR LOCATION.
RC TISSUE=Colon;
RX PubMed=12662155; DOI=10.1042/BJ20021914;
RA Qi S.Y., Riviere P.J., Trojnar J., Junien J.-L., Akineanya K.O.;
RT "Cloning and characterization of dipeptidyl peptidase 10, a new member
RT of an emerging subgroup of serine proteases.";
RL Biochem. J. 373:179-189(2003).
[3]
RN NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1), PARTIAL NUCLEOTIDE SEQUENCE
RN [MRNA] (ISOFORM 2), CATALYTIC ACTIVITY, BIOPHYSICO-CHEMICAL PROPERTIES,
RN TISSUE SPECIFICITY, AND SUBCELLULAR LOCATION.
RX PubMed=15245913; DOI=10.1016/j.bbexp.2004.03.010;
RA Ajami K., Abbott C.A., McCaughan G.W., Gorrell M.D.;
RT "Dipeptidyl peptidase 9 has two forms, a broad tissue distribution,

RT cytoplasmic localization and DPV-like peptidase activity.";
 RL Biochim. Biophys. Acta 1679:18-28(2004).
 RN [4].
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 1).
 RC TISSUE=Placenta, and Skin;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Scapellato M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Brownstein M.J., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Pahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bonfard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [5].
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 30-863 (ISOFORM 3),
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 272-863 (ISOFORM 2), AND
 RP PARTIAL NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 5).
 RC TISSUE=Glial tumor, Ovary, Spleen, and Trachea;
 RX PubMed=14702039; DOI=10.1038/ng1285;
 RA Oba T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
 RA Wakamatsu A., Hayashi K., Sato H., Negai K., Kimura K., Makita H.,
 RA Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
 RA Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,
 RA Nagahori A., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,
 RA Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,
 RA Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,
 RA Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M.,
 RA Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,
 RA Fujimori K., Tanai H., Kimata M., Watanabe M., Hiraoka S., Chiba Y.,
 RA Ishida S., Ono Y., Takiguchi S., Watanabe M., Yosida M., Hotuta T.,
 RA Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,
 RA Nomura Y., Togiyasu S., Komai F., Hara R., Takeuchi K., Arita M.,
 RA Inose N., Musashino K., Yuki H., Oshima A., Sasaki N., Aotsuka S.,
 RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohara T., Sano S.,
 RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
 RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
 RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
 RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
 RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
 RA Ono T., Yamada K., Fujii Y., Ozaki K., Hiroo M., Ohmori Y.,
 RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
 RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
 RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
 RA Togaishi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
 RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
 RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
 RA Nakai K., Yada T., Nakamura Y., Ohara O., Isoigai T., Sugano S.;
 RT "Complete sequencing and characterization of 21,243 full-length human
 RT cDNAs.";
 RL Nat. Genet. 36:40-45(2004).
 RN [6].
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 209-863 (ISOFORM 4), AND
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 298-863 (ISOFORM 2).
 RC TISSUE=Melanoma;
 RG The German cDNA consortium;
 RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Dipeptidyl peptidase that cleaves off N-terminal
 CC dipeptides from proteins having a Pro or Ala residue at position
 CC 2.
 CC -!- CATALYTIC ACTIVITY: Release of an N-terminal dipeptide, Xaa-Yaa-|-
 Zaa-, from a polypeptide, preferentially when Yaa is Pro, provided
 Zaa is neither Pro nor hydroxyproline.
 CC -!- ENZYME REGULATION: Inhibited by the serine proteinase inhibitor 4-
 CC (2-aminoethyl)benzenesulphonyl fluoride (AEBSF), and by di-
 CC isopropylfluorophosphate.
 CC -!- BIOPHYSICOCHEMICAL PROPERTIES:
 CC Kinetic parameters:
 CC KM=161 uM for Ala-Pro-AMC;
 CC KM=180 uM for Ala-Pro-AFC;
 CC pH dependence:
 CC Optimum pH is 7.5-8.5. Little activity below pH 6.5;
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic; cytosol.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=5;
 CC Name=1; Synonyms=Short;
 CC IsoId=Q86T12-1; Sequence=Displayed;
 CC Name=2; Synonyms=Long;
 CC IsoId=Q86T12-2; Sequence=VSP_013865;
 CC Note=Incomplete sequence;
 CC Name=3;
 CC IsoId=Q86T12-3; Sequence=VSP_013867, VSP_013868;
 CC Name=4;
 CC IsoId=Q86T12-4; Sequence=VSP_013869;
 CC Name=5;
 CC IsoId=Q86T12-5; Sequence=VSP_013865, VSP_013866;
 CC Note=Incomplete sequence;
 CC -!- TISSUE SPECIFICITY: Ubiquitously expressed, with highest levels in
 CC liver, heart and muscle, and lowest levels in brain.
 CC -!- SIMILARITY: Belongs to the peptidase S9B family. DPPIV subfamily.
 CC -!- CAUTION: Ref.6 (CAD39039) sequence differs from that shown due to
 CC frameshifts in positions 432 and 460.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 CC EMBL; AF452102; AAL47179.1; -; mRNA.
 CC EMBL; AY172660; AAC17262.1; -; mRNA.
 CC EMBL; AF542510; AAC073880.2; -; mRNA.
 CC EMBL; AY374518; AAQ83119.1; -; mRNA.
 CC EMBL; BC000970; AAH00970.1; -; mRNA.
 CC EMBL; BC037948; AAH37948.1; -; mRNA.
 CC EMBL; AK054656; BAB70784.1; -; mRNA.
 CC EMBL; AK075030; BAC1362.1; -; mRNA.
 CC EMBL; AK131499; BAC85150.1; -; mRNA.
 CC EMBL; AK131499; BAD18643.1; ALT INIT; mRNA.
 CC EMBL; AL834376; CAD39039.3; ALT_FRAME; mRNA.
 CC EMBL; CR827380; CAH10477.1; -; mRNA.
 CC MEROPS; S09.019; -.
 CC EnsEmbl; ENSG00000142002; Homo sapiens.
 CC HGNC; HGNC:18648; DPP9.
 CC MIM; 608258; -.
 CC InterPro; IPR001375; Peptidase_S9.
 CC InterPro; IPR002469; Peptidase_S9B.
 CC InterPro; IPR000379; Ser.estr.
 CC Pfam; PF00930; DPPIV_N; 1.
 CC Pfam; PF00326; Peptidase_S9; 1.
 CC Alternative splicing; Amino-peptidase; Hydrolase; Protease;
 KW Serine protease.
 FT ACT_SITE 730 730 Charge relay system (By similarity).
 FT ACT_SITE 808 808 Charge relay system (By similarity).
 FT ACT_SITE 840 840 Charge relay system (By similarity).
 FT VARSPLIC 1 1 M -> LSRVPCVRGCRPPPLPPLPGSOSRAWSDRRLAPL
 FT RLMRKVKLRLDKENTGWSRFSLSNESAER (in
 FT isoform 2 and isoform 5).
 FT /FTId=VSP_013865.
 FT VARSPLIC 3 139 Missing (in isoform 5).
 FT /FTId=VSP_013866.
 FT VARSPLIC 650 674 QLVNNSFGIKYLRINTLASLGYAV -> SAHLLPRPPPHH
 FT PEDSPSLKCL (in isoform 3).
 FT

[illegible]

Tue Apr 18 08:18:35 2006

```

Db      386 AIAGAPVTVMAYDTGYTERMDVPENNOHQYEAGSVALHVEKLEPNERNLLILHGFIDE 445
Qy      819 NVHFAHTSLLSLVRAKGPYDL-----OYPOERHSIRVPESGEHYELHLLHYIQ 869
Db      446 NVHFFHTNFLVSQLIRAKGPVQLQVALPPVSPQIYPNERNHSIRCPESGEHYEVLTHFLQ 505
Qy      870 ENL 872
Db      506 EYL 508

RESULT 8
ID Q7QBKI_ANOGA PRELIMINARY; PRT; 886 AA.
AC Q7QBKI;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ENSANGP0000016526 (Fragment).
GN ORFNames=ENSANGG0000014037;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Culicidae;
OC Anophelinae; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RT "Anopheles gambiae re-annotation.";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AA0801008879; EAA08416.2; -; Genomic_DNA.
DR MEKOPS; S09.016; -; Catalytic activity; IEA.
DR GO; GO:0003824; F:kinetic-type peptidase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001375; Peptidase_S9.
DR InterPro; IPR002469; Peptidase_S9B.
DR InterPro; IPR000379; Ser_estrs.
DR Pfam; PF00930; DPPIV_N; 1.
DR Pfam; PF00326; Peptidase_S9; 1.
FT NON_TER
FT NON_TER
SQ SEQUENCE 886 AA; 100440 MW; 3C284605CAA57DB4 CRC64;

Query Match 34.9%; Score 1641; DB 2; Length 886;
Best Local Similarity 39.2%; Pred. No. 1.8e-110;
Matches 361; Conservative 148; Mismatches 287; Indels 124; Gaps 20;

Qy      41 SWSLKKLLADTRKYHYGMMAKA-PHDFMFVKNRDPDGFHSDRIYLLAMSGENRNT-LF 98
Db      2 SWSLKCIVNTRQMATTAASFPMSVNF--RTLSDG--RTRVYFLSPPPANGWDTILF 57
Qy      99 YSEIPKTIINRAAVL-----LSWKPLLD 121
Db      58 YADVPPAGAKRSQLLPALDANELEADGSDGDEDAEGERSSGAKRKHLPWQOLLE 117
Qy      122 LFOATLDYGVSRREELRLRRKIGTVGTASYDYHQSGTFLFQAGSGIYHKDGPQGF 181
Db      118 SVLGHLSLTS-NSREVQMLERKSLTWGTSVELHKAAGKIVFPACNTLYQCLD---TG 173
Qy      182 TQOPLRPNLVETSCPNIRMDPKLPADPDWIAPIHSDNIWISNIVTREERLTYVHNELA 241
Db      174 EENPLFSELRIQRAAIDPQICPQNSDLVAFVNCNGDIWVHVTHSGHSERLTYAHGRR 233

```

```

Qy      242 NNEEDARSAGVATVLOEEFDRIYSGYWCPCKAETTPSGGKILRLIYEENDESEVEIHHVT 301
Db      234 SFAEDPLTAGVSPSYVMQEEFESRIQGFWMQPE-----SHDEVYRIYVEEVDESDSLTYFP 288
Qy      302 SMLETRADSPRYPKTGTPANKPKVTFKMSHIMIDASGRIDVIDKELIOPFEILPEGVY 361
Db      289 SSQSAGRDYEBYRFRAGTPTNAKSLKLVQFRUSENLRTDVCIKELQCLTFAPFWLEY 348
Qy      362 IARAGWTEGKYAWSILLDRSQTLQVLISPLFI-----PVEDDMERORLIE 411
Db      349 IVRVGWTPDSRYVWAQLDRPQORLELVLLPVPNFCFIYSSSSSHHPKSSSGWSRPLDK 408
Qy      412 SVPDSVTEL-IYBETTDIMINIHDPHVPQSHSEIEPIFASCEKTCFRHLYKITSIL 470
Db      409 S---TTRPLQVIYTTETSSWVNVHDVLQ-FVELSEQEVTFMASE-ESGFRHLILYVSSL 463
Qy      471 KESKYKRSRSGGLPAPSPDFKCP-----IKEEIAITSGEWVNLGRHSGNSIOVDEV 518
Db      464 SPNEV--SSGGVGA-TDHSLSMACIGSTLVARIVQKVTLTGDDWEVLGR--NVMYDRA 517
Qy      519 RRLVVFEGTKDSPLHHLVYVYVNPGEVTRLTDRGYSHSCCISOHCDPFIKYSNQKMP 578
Db      518 RQLVYFMGLRETPELEKHLVYVSLAQPNQLRLTMPGYSFTVEFNDDCTTLFLOTYCNSTL 577
Qy      579 HCVSLYKL---SSPEDDPTCKTKFWMATILDSAGPLDPTP----- 616
Db      578 PSWELVRIAHDSNTANGNC-----SHGTPPTPTIDALRLCSVGYLTEGGPSE 626
Qy      617 -----PEIPESTETGFTLYGMLYKPHDLQPKKYPTVLFYIGPQVQLVNNRFGVKYF 671
Db      627 NTOYNPSIHSPOISSGDVLYAMVFKPHFMLGVKYPTVLNVYGGPEVQTVSNTFKGMROL 686
Qy      672 RLNTLASLQYVVVIDNRGSGHGLKKEGAPKYKMGQIEIDDOVEGLQYLASRYDFDILD 731
Db      687 RMELASQGYCVICVDSRGRHGVPEFSYICRMGTVELSDQVEVLRLADQLGVIDMD 746
Qy      732 RVGIHWSYGGYLSLMALMQRSDTPRVAIAGAPVTLFIYDVTGYTERYMGHPQDNBOGY 791
Db      747 RVAIHWMSYGGYLSLMLGVQVPEIFKYSIAGAPVTSWEYDYGTERYMDLPDSNRSGYA 806
Qy      792 LGSVAQMAKFPSEPNRLLHGLFDENVHFAHTSILLSVLRAKGYDQLIYPOERHSI 851
Db      807 AGSVLNYIQKFPDENRLLIHLGLDENVHFAHTSILLSVLRAKGYDQLIYPOERHSI 856
Qy      852 RVPESGEYELHLLHYLQEN 871
Db      867 RNLEASKHYETKLLSLFQNH 886

RESULT 9
Q9VC20_DROME
ID Q9VC20_DROME PRELIMINARY; PRT; 1053 AA.
AC Q9VC20_Q8IIH07;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE CG3744-PA, isoform A (CG3744-pc, isoform c) (Ld33755p).
GN ORFNames=CG3744;
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP NUCLEOTIDE SEQUENCE
RX MEDLINE=20196000; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.R., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Stoughton R.C., Rogers J., H.C. Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.C., Holt R.A., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-U., Andrews-Pfannkoch C., Baldwin D.,

```


RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA Posler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong P., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merklov G., Milshina N.V., Mobarry C., Morris J., Woshref A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Sinden-Kianos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Yel J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong P.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
RT "Finishing a whole-genome shotgun: release 3 of the Drosophila
RT melanogaster euchromatic genome sequence."
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RL [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426065; PubMed=12537568;
RA Ceiniker S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R., Lavery T., Muzny D.M., Nelson C.R.,
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.,
RT "Finishing a whole-genome shotgun: release 3 of the Drosophila
RT melanogaster euchromatic genome sequence."
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RL [3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminiker J.S., Bergman C.M., Krommiller B., Carlson J.W., Svirskas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Ceiniker S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
RT a genomics perspective."
RL Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20(2002).
RL [4]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Ceiniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review."
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
RL [5]
RP NUCLEOTIDE SEQUENCE.
RX Berkeley Drosophila Genome Project;
RA Ceiniker S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,
RA Hoskins R., Stapleton M., Pacleb J., Park S., Svirskas R., Smith E.,
RA Yu C., Rubin G.;
RT "Drosophila melanogaster release 4 sequence."
RT Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

RNA NUCLEOTIDE SEQUENCE.
RPFlyBase;
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
RL [7]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Berkley;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Ceiniker S.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
CC -!- INTERACTION:
CC QGVZRO:CG12016; NBExp=1; IntAct=EBI-105926, EBI-152819;
DR EMBL; AE003749; AAF5357.2; -; Genomic_DNA.
DR EMBL; BT001499; AAN71254.1; -; mRNA.
DR INACT; Q9VC20; -;
DR MEROPS; S09.016; -;
DR EMBL; FBgn0039240; CG3744.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0004274; F:dipeptidyl-peptidase IV activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001375; Peptidase S9.
DR InterPro; IPR002469; Peptidase S9B.
DR InterPro; IPR003379; Ser. esters.
DR Pfam; PF00930; DPPIV_N; I.
DR Pfam; PF00326; Peptidase S9; 1.
DR SEQUENCE 1053 AA; 118053 MW; C94AA663AB464577 CRC64;
QY Query Match 34.0%; Score 1596; DB 2; Length 1053;
Db Best Local Similarity 37.9%; Pred. No. 4.6e-107;
QY Matches 369; Conservative 159; Mismatches 307; Indels 138; Gaps 27;
Db 2 AAAMETQL-----GVEIFETAD-CEENIESQDRPKLEPFV-----ERYSWSLK 46
112 AASVITNHLSSPTSGTPPHGLDVGDEDCDEDEDDPDVNDGHIAPTNNKSWAENK 171
QY 47 KLLADTRKHYGMWAKPHDFMFKRNDPDGPHSDRIYVLAMSGENRENTFYSPKPTI 106
Db 172 QVQVEIRKMCNLSWVPTVQF--RHLSDG--RARCYLEGTGTPPSWETILLFADINLTQ 227
QY 107 NRAAVLM-----LSWKPLLDL-FOAT 126
Db 228 SEEQQLLVQRLEGIASDEWSPWNAGSTSSGHQPAFLFNSLPRLPWSPLLOQPIQSS 287
QY 127 LDYG-----MYGEEELLRKRIGTVGIASVDYHOGSGTFLQAGSGIYHVKGQPGQ 180
Db 288 GSGSGSGSASPYAREYQLQERKRLSTWGITSYELHKPSGKLVFPFNDLYQCLD---TG 344
QY 181 FTQQLPRLNVLVETSCPN-IRWDPKLCPADPDWTFATHSNDIWSINLVTRERRLTYVHNE 239
Db 345 YNSGLLPPTQRLT-CPQWALTDPQICQNSDMIAIYISDCDLFTVHTLSGHEKRLTYSTG 403
QY 240 LANNEEDARGAGNATFVLQEEFDYRGYGMWCPKAETTPSGGKILRLIYEENDESVEIHI 299
Db 404 RHSVYDDALSGVPSYVMQEEFSRYQGFWMQPH-----SNDGIYRIVVEEVESESVVT 458
QY 300 VTSPLMELTRRADSPRYKPTGTANKPVTFKMSIEMIDAEGRIIDVDKELIQPEIL--FE 357
Db 459 FFSSTAMHGRVDEYRFPRTGSPNAKSLKLVQVPLNEALQVSEIAIKOL--PYSLLAVFS 516
QY 358 GVEIYARAGWTPECKYAWSTLLDRSOTRLQIVLISPELFIPEVDDWMEORLIESVPD-- 415
Db 517 WLEIYVGVTFDQAYVWVGLDKKQORLDVILIPDNTF--CESYSSQVSTFTDSIGDS 574
QY 416 -----SVTPL-IYVETTDIWINIHDIFHPVFPQSHEEIEFIFASECKTGFRHLYKIT 467
Db 575 WRSLSYRTITPLQVIYTERSDSMINVDMLH-FIDLTTSTVTFMWASE-ETGFRHLYLVT 632

```
QY 468 SILKESKYSRSGGLPAP-----SDFKCPKEIEIATSGEWEVLGRHGSNTQV 515
DQ 633 ASLLLSQ-----ANGQPDGSGVGAQPSFVDSALQPRILNKVATSGEWEVLAR---NLWV 685
QY 516 DEVRRLVYFEGTKQSPLEHLYVYVYNPGVEVTRLTDRGYSHSC-----CISQH 564
DQ 686 DRANKLVYFVGLRDTPEKHLVYVYSLRPEPHIRLLTEPGSYLVLEFDDHFNDLFFIFSQQ 745
QY 565 CDFFIKYSKN-QKNPHCVSLYKLSGSPDDPTCKTFEAWILDSAG-----PLPDYTP 616
DQ 746 CKMLLVYCNIRLPSCVKM-----RVNQTCSNGVNGIQISLVGLYHGGKPEQVYC- 798
QY 617 PEIFSFESTGTGLVGLYKPHDLPQKKYPTVLFIYGGPQVQLVNNRFGVKYKRLNLT 676
DQ 799 PQIFSPQLPSGDIYVAVFKPHNFELGVKYPTVLNVYGGPEVQTVANTFKGKHQLRMML 858
QY 677 ASLGVYVVVINDRSGCHRLKFGAFKVKMGQLEIDDOVEGLQYLASRYDFIDLDRVGTH 736
DQ 859 AAGQYCVICIDBSGRHGRKGFESHIRGRMGQVELTDOVALRSLDQLGYIDMRVAIH 918
QY 737 GHSYGGYLSMALMQRSDFRVAITAGAPVTLWTFYDGTGYRYVMGHDPDQNEGYVLGVA 796
DQ 919 GMSYGGYLSMGLVQYPIKFAITAGAPVTWNEYIDGTGYRYVMGHDPDQNEAGYSAGSVL 978
QY 797 MQAEKPPSPENRLLHGLFENVHFPAHTSILLSFLVRACKPYDLQIYQERHSYRPES 856
DQ 979 EYVNFPPEDKRLHGLIDENVHFCHTSLRSLANKANKPYEVHLFFEERHSLNLES 1038
QY 857 GEHYELHLLHYLQ 869
DQ 1039 NKNYETKLLSLFQ 1051

RESULT 10
ID Q9VC19 DROME PRELIMINARY; PRT; 1113 AA.
AC Q9VC19;
AT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE CG3744-PB, isoform B.
GN Names=CG3744; ORFNames=CG3744;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Phyllostoma; Drosophilidae; Drosophila.
OX NCBI_Taxids=7227;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champagne M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballou R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brothier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA de Pablo B., Delcher A., Deng Z., Davenport L.B., Davies P.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Fertiera S., Fleischmann W.,
RA Flesler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glöck A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Hartin D., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Jallali M., Kalush P., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
```

```
RA Merklov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson K.A., Nixon K., Nusskern D.R., Paclebb J.M., M.G.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Fur V., Reese H.,
RA Reinart K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun B.,
RA Svirkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zhou L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champagne M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,
RA Paclebb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirkas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstein G., Scher S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun, release 3 of the Drosophila
RT melanogaster euchromatic genome sequence.";
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Krommiller B., Carlson J.W., Svirkas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
RT a genomics perspective.";
RL Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20(2002).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review.";
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RX Berkeley Drosophila Genome Project;
RA Celniker S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,
RA Hoskins R., Stapleton M., Paclebb J., Park S., Svirkas R., Smith E.,
RA Yu C., Rubin G.;
RT "Drosophila melanogaster release 4 sequence.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [6]
RP NUCLEOTIDE SEQUENCE.
RX FlyBase;
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003749; AAF56356.2; -; Genomic_DNA.
DR MEROPS; S09.016; -.
DR Ensembl; CG3744; Drosophila melanogaster.
DR FlyBase; FBgn0039240; CG3744.
DR GO; GO:0003824; F: catalytic activity; IEA.
DR GO; GO:0008236; F: serine-type peptidase activity; IEA.
DR GO; GO:0006508; P: proteolysis and peptidolysis; IEA.
DR InterPro; IPR001375; Peptidase_S9.
DR InterPro; IPR002469; Peptidase_S9B.
DR InterPro; IPR000379; Ser. esters.
DR Pfam; PF00930; DPPIV_N; 1.
DR Pfam; PF00326; Peptidase_S9; 1.
```

```
SQ SEQUENCE 1113 AA; 123881 MW; 20857E3B212DF2E4 CRC64;

Query Match 34.0%; Score 1596; DB 2; Length 1113;
Best Local Similarity 37.9%; Pred. No. 5e-107;
Matches 369; Conservative 159; Mismatches 307; Indels 138; Gaps 27;

QY 2 AAAMETEQL-----GVEIFETAD-CEENTESQDRPKLEPFYV-----ERYVSOLK 46
DB 172 AASVITNHLSSPTSGTPPHGLVDDEGDDCEDEDDPVDNDGHIAAPTPNKNWAENK 231
QY 47 KLLADTKYHGMKAPADPMFKVRNPDGPHSDRIYIYLAMSGENRENTLFYSEIPKTI 106
DB 232 QVQOEIRKMCNLSMWPTNVQF--RHLSDG--RARCFLGTPTQSQWETILLFADINLTQ 287
QY 107 NRAAVLM-----LSWKPLDL-FOAT 126
DB 288 SEEQQLLVQRLGASDEWSPPTWAGSPPTSSGHQPAFLFNSLPRLPWSPLLOQPIQS 347
QY 127 LDYG-----MYGREBELLRKRIGTVGIASVDYHOGSGTFLFOAGSGIYHVKGQPG 180
DB 348 GSGGSGSASPYAREYQLQERKLSWTGITSYELHKPSGKLVPCFNDLYQCLD---TG 404
QY 181 FTQOPLRPNLVETSCN-TRMDPKLCPADPDWIAIHNSDIWISNIVTREERRLTYVHNE 239
DB 405 YNSGLLPFTQLRT-CPQWALDPOICPQNSDMIAYISDCDLFVTHLSGHEKRLTYTSTG 463
QY 240 LANNEEDARAGVATVLOEEDFRIYGYWCPKATETSSGKILARILYENDESEVPIIH 299
DB 464 RHSVYDDALSAGVPSYVMQEFESRYQGFWMQPH-----SNDGIYRIYVEEVESEVYT 518
QY 300 VTSPLMLETTRADSFYRPKGTANPKVTFPMWSEIMIDAEGRIIDVIDKELIOPREIL--PE 357
DB 519 FPSSTAMHGRVDEYRFPRTGSPNASKLKLQVFLNEALQVSEIAKDL--PYSLLA VFS 576
QY 358 GVEIARAGTTPGKAWSTLLDRSOTRLQIVLISPELFIPEVDDVMERQRLIESVDP-- 415
DB 577 WLEYIVRVGWTPOAKYVWVQGLDRKQOORLDVILIDLPNF--CESYSSQVSTPTDSIGDHS 634
QY 416 -----SVTPL-IIEEETDIWINIHDIFHFVPOSHEEIEFIPASECKTGFRHLYKIT 467
DB 635 WRSLSYSTIITPLQVITERSDSWINVHMLH-FLDITSTVTFWLWASE-ETGFRHLYVT 692
QY 468 SILKESYKRSSGGLPAP-----SDFKCPKEEIAITSGEWEVLGRGSGNIQV 515
DB 693 ASLLLSQ-----ANGQPDGSGVGAQPSFVDSLSALQPRILNKVALTSGEWEVLAR--NLW 745
QY 516 DEVRALVYFGTKDPSLEHLLHYVSVYVNGEVTRLTDRGYSHSC-----CLSQH 564
DB 746 DKANKLVYFVGLRDTPLEKLYVVSLSRPERHILLTEPGSYLVLEFDDHFDNLPFIQSQ 805
QY 565 CDFIFISKYSN-QKNPHCVSLYKLSSEDDPTCKTEFWATILDSAG-----PLPDYTP 616
DB 806 CKMLLYVCNIQLPSCKNV-----RVNQTCNSGVNGIQISLVGYLHEGKPEPQYC- 858
QY 617 PEIYFSTFTCTGLYMLYKPHDLQPKKYPTVLFIYGGPQVQVNNRFKGVYFRNLTL 676
DB 859 PQIFSPOLPSGDIYAMVFKPHNFGVYPTVNLVYGGPEVQTVNNTFKGKHQLRHML 918
QY 677 ASLGVVVVINDGSGCHRGKFGAFKYMGOIEIDDOVEGLOYLASRYDFIDLDVIGIH 736
DB 919 AAQGYVICIDSGSRHRGKRFESHIRGRMGQVELFDQVDAULRSQDGLYIDMDRVAIH 978
QY 737 GWSYGGYLSMALMQRSDIFRVAIAGAPVTLTIFDYTGTYTERYMGHPQNEQGYLGSVA 796
DB 979 GWSYGGYLSMLVQYFKIYKVAIAGAPVNTWEYDYGTYTERYMDQPNQNEAGYSAGSVL 1038
QY 797 MQAKFPSEPNRLLLHGLFDENVHFAHTSILLSPLVRAGKYDQIYVQERHSIRVPES 856
DB 1039 EYVNSFPEDKRLLIHGLIDENVHFCHTSRLISALNANKPKYEVHLFPEERHSLRNLES 1098
QY 857 GEHYELHLLHYLQ 869
DB 1099 NKNYETKLSPLQ 1111
```

RESULT 11

```
Q5TTK8 ANOGA PRELIMINARY; PRT; 740 AA.
AC Q5TTK8;
DT 01-FEB-2005 (T-EMBLrel. 29, Created)
DT 01-FEB-2005 (T-EMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (T-EMBLrel. 29, Last annotation update)
DE ENSANGP0000026132 (Fragment).
GN ORFNames=ENSANGG0000014037;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Culicidae;
OC Anophelinae; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG "The Anopheles gambiae Sequence Committee;
RT "Anopheles gambiae re-annotation."
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AA0801008679; EAL40717.1; -; Genomic DNA.
DR GO; GO:0003824; F: catalytic activity; IEA.
DR GO; GO:0008236; F: serine-type peptidase activity; IEA.
DR GO; GO:0006508; P: proteolysis and peptidolysis; IEA.
DR InterPro; IPR001375; Peptidase S9.
DR InterPro; IPR002469; Peptidase S9B.
DR InterPro; IPR000379; Ser esters.
DR Pfam; PF00930; DPPIV N, I.
DR Pfam; PF00326; Peptidase_S9; 1.
FT NON TER 1
SQ SEQUENCE 740 AA; 84151 MW; D59F131BBA095CCB CRC64;

Query Match 33.3%; Score 1567; DB 2; Length 740;
Best Local Similarity 42.3%; Pred. No. 3.6e-105;
Matches 330; Conservative 129; Mismatches 241; Indels 80; Gaps 15;

QY 133 SREBELLRKRIGTVGIASVDYHOGSGTFLFOAGSGIYHVKGQPGTQOPLRPNLVE 192
DB 1 SREVQLMERKRUSIGWITSYELHKASGKIVFPACNTLYQCLD---TGYEENPLFPSEL 57
QY 193 TSCPNIRMDPKLCPADPDWIAIHNSDIWISNIVTREERRLTYVHNELANMEEDARSAGV 252
DB 58 ILQRAAALDPOICQNSDLVAFVCGDIWVHTSHGSHSELTAYHDGRRSPAEDPLTAGV 117
QY 253 ATFLVQBEFDRYGYWVCPKAEETTPSGK-----ILRILYENDESEVEIILHVTSP 303
DB 118 PSYVQBEFSRYQGFWMQPE-----SHGKFDMDVILDEVYRIVYEEVDSVLTFFSS 172
QY 304 MLETRRADSPRYKGTANPKVTFKMSIEMIADGRIIDVIDKELIQLPEILFEGVEYIA 363
DB 173 QSAGRDYEEYRFPAGTPNAKSKLKVQFRLSENRLITDVCIKELQCLPTFAPPWLEYIV 232
QY 364 RAGWTPGKYAWSTLLDRSOTRLQIVLISPELFIPEVDDVMERQRLIESVP-----DS 416
DB 233 RVGTPDSRYVWVAQLDRPQORLELVLLPVDNFCBIYS-----SSRPPAGARRKST 284
QY 417 VTPLE-IIEEETDIWINIHDIFHFVPOSHEEIEFIPASECKTGFRHLYKITSLKESKY 475
DB 285 TRPLQVITYTETSSWNVNVDVLQ-FVELSQEVTFLWASE-ESGFRHLYLVT----- 335
QY 476 KSSGGLPAPSDFKCP-----IKEEIAITSGEWEVLGRHSGNIQVDEVRRLVY 523
DB 336 KVSSGGVGA-TDHSLSPMACIGTLVARIQVKTLTGQDWEVLGR---NWYDRAQLVY 391
```



```
DR GO:0003824; F:catalytic activity; IEA.
DR GO:0008236; F:serine-type peptidase activity; IEA.
DR GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001375; Peptidase_S9.
DR InterPro; IPR002469; Peptidase_S9B.
DR InterPro; IPR00379; Ser_estr.
DR Pfam; PF00930; DPPIV_N; I.
DR Pfam; PF00326; Peptidase_S9; 1.
FT NON_TER 557
FT NON_TER 557
SQ SEQUENCE 557 AA; 62685 MW; 08AAAA944ACD0808 CRC64;

Query Match 24.38; Score 1142.5; DB 2; Length 557;
Best Local Similarity 43.38; Pred. No. 2.2e-74;
Matches 253; Conservative 84; Mismatches 198; Indels 49; Gaps 12;

QY 284 RILYEENDESEVEIHTVTSPLMSTRADSFYRPTKGTANPKVTFKMSIMIDAGRIIDV 343
Db 1 RIVYEENDESDVSLYFPSSQSGRDYBEYRFRAGTPNNAKSLKLVQFRLSENRLTIDV 60
QY 344 IDEKLIQFELFEGVEIYARAGTPEGYAWSLILDRSQTRLQIQLISPELIPVEDDV 403
Db 61 CIKELQCLPLFAFPWLEYIVRGVCTPDSRYVWAQLLDRPQORLELVLPVDFCSTA--A 118
QY 404 MERQRLIESVPDSVTPLIIVEETTDIWINIHDI FHVFPQSHHEEIEFIFASECTGPRHL 463
Db 119 RRRRPTAGRPAGGAMGPGIYVTSSTSSWNVHVDLQ--FVELSDEY-----TTNSHP 168
QY 464 YKITSILKESKY--KRSSGGLPAPDFKCP-----IKKEIAITSGEVEVLGRH 509
Db 169 TAPKQOREGAGHTEVSSGAGA--TDHSLPSMACIGSTLVARIVQKVLTLGGDWELGR-- 226
QY 510 GSNLOVREVLVYFEGTKDSPLEHLYVSYVNPCEVTRLTDRGYSHSCISQHCDFP-- 568
Db 227 --NWDYDRVLQVYFMGLRETPLEKLYVYSLAQNQLRLTTPGYSFTV---BFNDLVR 281
QY 569 ISKYSNOKNCHVSLYKLSPEDDPTCKTKFEWATILDSAGLPDYP-----PFIIS 621
Db 282 IAHDSNTANGGCS-----HGTPPTP-----PIDALRLCSGVLTGGGSENVTQNPISHS 333
QY 622 FESTTGTFLYMLKPHDLQPKKYPVTLFIYGGPOVLVNNRFGVKYFRLNTLASLGY 681
Db 334 PQISGGDLVAMFKPHFNLGVKYPTVNLVYGGPEVQVSTNFKGQRLRWMLASQGY 393
QY 682 VVVVIDNRGSGHRLKPEGAFKYMGOIEIDQVEGLQVLAIRYDFIDLDRVGHWSYG 741
Db 394 CVICVDSRSGRHRGVEFESYIRRMGTVELSDQVEVLRIADQLGYIDMDRAVIGHWSY 453
QY 742 GYLSLMALMQRSDFIRVAIAGAPVTLWIFDYGTYERYMGHPDQNEQGYILGSAVMAQEK 801
Db 454 GYLSLMLGVQYPEIFKYSIAGAPVTSWEYDYGTYERYMGHPDQNEQGYILGSAVMAQEK 801
QY 802 FPSPENRLLHGLFDENVHFAHTSILLFLVRAGKPYDLQIYP 845
Db 514 FPDENRLLIITHGLIDENVHFAHTSILLFLVRAGKPYDLQIYP 557

RESULT 14
Q54U01.DICDI PRELIMINARY; PRT; 803 AA.
AC Q54U01;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=DD0205566;
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AX4;
RA Eichinger L., Pachebat J.A., Gloeckner G., Rajandream M.-A.,
```

QY 674 NTLASLGVVVVVIDNRGSGCHRLKPEGAFKYKMGQIEIDDQVEGLQYLASRYDF--IDLQ 731
Db 610 N-----FGFTQVMIDNVGSANRGLSEFESHIREKMGQVEIGDQVEGINYLGN-DIVSIDVN 664

QY 732 RVGIHGWSVGGYLSLMLMQRSDFRVAIAGAPVTWIFYDTGTYRYMGHPDQNEQGY 791
Db 665 RTAISGWSYGGYNSLMAISQRPDVFKIAVCGAPVSDWRLYNTGTYRYMNVDPQNDIDYK 724

QY 792 LGSVAMQAEKPESEPNRLLLHGLDENVHFHTSILLSFLVRAGPYDLQIYPOERHSI 851
Db 725 LGDTTHYS--FTEENRLLLIHGLQDENVHFSNTIYIIDLTKTKPKYILKTLPLNERHGV 782

QY 852 RVPESGEHYELHLHYQENL 872
Db 783 RNTDNRIYIGLFIINHLKLN 803

RESULT 15
O75868 HUMAN
ID O75868 HUMAN PRELIMINARY; PRT; 432 AA.
AC O75868;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
DE R33083_1 (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Lemerding J.E., McCready P.M., Skowronski E., Wiswanathan V.,
RA Burkhardt-Schultz K., Gordon L., Dias J., Ramirez M., Stilwagen S.,
RA Phan H., Velasco N., Do L., Regala W., Terry A., Garnes J.,
RA Danganan L., Erler A., Christensen M., Georgescu A., Avila J., Liu S.,
RA Attix C., Andrise T., Trankheim M., Amico-Keller G., Coefield J.,
RA Duarte S., Lucas S., Bruce R., Thomas P., Quan G., Kronmiller B.,
RA Arellano A., Saunders C., Ow D., Nolan M., Trong S., Kobayashi A.,
RA Olsen A.S., Carrao A.V.,
RA Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
RL EMBL; AC005783; AAC62840.1.; Genomic DNA.
DR Ensembl; ENSG00000142002; Homo sapiens.
DR InterPro; IPR002469; Peptidase_S98.
FT Pfam; PF00930; DPPIV_N; 1.
FT NON_TER 432
SQ SEQUENCE 432 AA; 48595 MW; 64E2B85BE0523A7E CRC64;

Query Match 22.3%; Score 1049; DB 2; Length 432;
Best Local Similarity 55.8%; Pred. No. 1e-67; Indels 8; Gaps 2;
Matches 189; Conservative 63; Mismatches 79;
QY 35 FYVERYSQKLLADTRKYHYGMAKAPHDFMFKRNDPDGPHSDRIYYLLAMSGENRE 94
Db 95 FVQKHSWDGLRSIIHGRKYSGLIYNKAPHDFQVQKTDSGPHSHRLYYLGMFYGSRE 154
QY 95 NTLFYSEIPKTIINRAAVLMSKPLDLDFQATLDYGMYSREELRLRKRKIGTVGIASYD 154
Db 155 NSLLYSEIPKKVRKEALLLSWKQMLDHFQATPHGYVSREELRLRKRKLGVEGITSYD 214
QY 155 YHOGSGTFLFOAGSGIYHVKGPGQF-----TQQLRPNLVETSCPNIRMDPKLCPA 207
Db 215 FHSESGFLFOAGSNLPHCRDGGKNGFMVSPGCVSPMKLEIKTCGSPRMDPKICPA 274
QY 208 DPDIATFHSNDIWSIVTREERRLTYYVHNELANNEEDARSAGVATVLOEEFDYSGY 267
Db 275 DPAPFSFINNSDLVANIETGEERRLTFCQGLSNVLDPPKSGAVATFVIOEEFDRFTGY 334
QY 268 WWCPCAEPTPGSKILRIIYEENDESEVEIHVTSPLMTRADSFRYPKGTGTANPKVTF 327
Db 335 WWCPTA-SWEEGLKTLRIIYEVEDESEVEIHVPSPALERKTDSDYRYPRTGSKNPKIAL 393

Search completed: April 14, 2006, 12:50:08
Job time : 246 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 14, 2006, 12:50:22 ; Search time 49 Seconds
(without alignments)
1488.162 Million cell updates/sec

Title: US-10-825-632-1
Perfect score: 4700
Sequence: 1 MAAMETEQLGVIEFETADC.....HLLHYLQNLGSRIFAALKVI 882

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2_6/ptodata/1/1aa/5 COMB.pep: *
2: /cgn2_6/ptodata/1/1aa/6 COMB.pep: *
3: /cgn2_6/ptodata/1/1aa/H COMB.pep: *
4: /cgn2_6/ptodata/1/1aa/PCTUS COMB.pep: *
5: /cgn2_6/ptodata/1/1aa/RE COMB.pep: *
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4700	100.0	882	2	US-09-976-674-1
2	4700	100.0	882	2	US-10-070-464-1
3	3504	74.6	658	2	US-09-976-674-19
4	3504	74.6	661	2	US-09-976-674-11
5	3504	74.6	690	2	US-09-976-674-7
6	3236	68.9	613	2	US-09-976-674-21
7	2870	61.1	863	2	US-09-976-674-3
8	2870	61.1	892	2	US-09-976-674-23
9	2870	61.1	892	2	US-09-976-674-27
10	2820.5	60.0	879	2	US-09-976-674-33
11	2820.5	60.0	879	2	US-09-976-674-35
12	2422	51.5	465	2	US-10-070-464-5
13	2406	51.2	832	2	US-09-976-674-29
14	2406	51.2	832	2	US-09-976-674-31
15	2356.5	50.1	819	2	US-09-976-674-37
16	2356.5	50.1	819	2	US-09-976-674-39
17	1836.5	39.1	360	2	US-10-070-464-7
18	1808	38.5	358	2	US-09-976-674-13
19	1645.5	35.0	310	2	US-09-794-236-4
20	1645.5	35.0	310	2	US-10-070-464-3
21	1391	29.6	518	2	US-09-976-674-25
22	1278	27.2	241	2	US-09-976-674-9
23	1007.5	21.4	194	2	US-09-976-674-17
24	719	15.3	981	2	US-09-902-540-16812
25	616	13.1	710	2	US-09-518-550-28
26	582.5	12.4	771	2	US-09-462-284-2
27	582.5	12.4	771	2	US-09-079-592-2

28	572.5	12.2	723	2	US-09-518-550-29	Sequence 29, Appl
29	572	12.2	732	2	US-09-518-550-30	Sequence 30, Appl
30	569.5	12.1	676	2	US-09-518-550-42	Sequence 42, Appl
31	529	11.3	766	2	US-10-002-593-6	Sequence 6, Appl
32	529	11.3	766	2	US-09-949-016-6146	Sequence 6146, Ap
33	529	11.3	766	2	US-10-423-714-6	Sequence 6, Appl
34	525	11.2	755	4	PCT-US93-07923-3	Sequence 3, Appl
35	525	11.2	759	4	PCT-US93-07923-2	Sequence 2, Appl
36	525	11.2	766	1	US-08-230-491A-3	Sequence 3, Appl
37	525	11.2	766	1	US-08-619-280A-3	Sequence 3, Appl
38	525	11.2	766	1	US-08-940-391-3	Sequence 1, Appl
39	525	11.2	766	2	US-09-794-236-1	Sequence 3, Appl
40	525	11.2	766	2	US-09-263-606-3	Sequence 10450, A
41	525	11.2	775	2	US-09-949-016-10450	Sequence 27, Appl
42	519	11.0	766	2	US-09-518-550-27	Sequence 11421, A
43	488.5	10.4	737	2	US-09-902-540-11421	Sequence 3, Appl
44	484.5	10.3	818	2	US-09-462-845-3	Sequence 3, Appl
45	484.5	10.3	818	2	US-10-402-312-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-09-976-674-1
; Sequence 1, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976, 674
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-1

Query Match	100.0%;	Score	4700;	DB 2;	Length	882;
Best Local Similarity	100.0%;	Pred. No.	0;			
Matches	882;	Conservative	0;	Mismatches	0;	Indels
						Gaps
Qy	1	MAAMETEQLGVIEFETADCENIESODRPKLEPFVVERYSWSQLKLADTRKYHGYMM	60			
Db	1	MAAMETEQLGVIEFETADCENIESODRPKLEPFVVERYSWSQLKLADTRKYHGYMM	60			
Qy	61	AKAPDPMFVKRNDPDPHSDRIYILAMSGENRENTLFYSEIPKTIINRAAVMLSWKPL	120			
Db	61	AKAPDPMFVKRNDPDPHSDRIYILAMSGENRENTLFYSEIPKTIINRAAVMLSWKPL	120			
Qy	121	DLFQATLDYGMYSREELLRRKRIGTVGIASDYHQSGTFLFOAGSGIYHVXDKGPGQ	180			
Db	121	DLFQATLDYGMYSREELLRRKRIGTVGIASDYHQSGTFLFOAGSGIYHVXDKGPGQ	180			
Qy	181	FTQOPLPNLVETSCNIRMDPKLCPADPDWIAFTHSNDIMISNIVTREERLTVVHNE	240			
Db	181	FTQOPLPNLVETSCNIRMDPKLCPADPDWIAFTHSNDIMISNIVTREERLTVVHNE	240			
Qy	241	ANMEEDARSAGVATFVLOEEDFRYSYWCPCAETTPSGGKILRLIYENDESEVEI	300			
Db	241	ANMEEDARSAGVATFVLOEEDFRYSYWCPCAETTPSGGKILRLIYENDESEVEI	300			
Qy	301	TSPMLETRRADSFYRPTGTANPKVTFKMSIMDAEGRIIDVIDKELIQPFELFEGVE	360			
Db	301	TSPMLETRRADSFYRPTGTANPKVTFKMSIMDAEGRIIDVIDKELIQPFELFEGVE	360			

QY	361	YIARAGWTPEGKYAWSILLDRSOTRLQIVLISPELFI PVEDDVMERQRLIESVPDSVTPL	420
Db	361	YIARAGWTPEGKYAWSILLDRSOTRLQIVLISPELFI PVEDDVMERQRLIESVPDSVTPL	420
QY	421	IYVEETTDIWINIHDIHVFPQSHHEEIEFIFASECKTGFRHLYKITTSILKESKYKRSG	480
Db	421	IYVEETTDIWINIHDIHVFPQSHHEEIEFIFASECKTGFRHLYKITTSILKESKYKRSG	480
QY	481	GLPAPSDPKCPKEBIAITSGEVEVLGRHGSNIQVDEVRLLVYFEGTKDSPLEHLLYVVS	540
Db	481	GLPAPSDPKCPKEBIAITSGEVEVLGRHGSNIQVDEVRLLVYFEGTKDSPLEHLLYVVS	540
QY	541	YINPGEVTRLDRGSHSCCISQHCDFFI SKYSNOKNPHCVSLYKLSSEDDPTCKTKEF	600
Db	541	YINPGEVTRLDRGSHSCCISQHCDFFI SKYSNOKNPHCVSLYKLSSEDDPTCKTKEF	600
QY	601	WATILDSAGPLPDYTPPEIFSFESTTGTLYGMLYKPHDLQPKKYPTVLFIYGGPQVQL	660
Db	601	WATILDSAGPLPDYTPPEIFSFESTTGTLYGMLYKPHDLQPKKYPTVLFIYGGPQVQL	660
QY	661	VNNRFGVKYFRNLNTLASLGYVVVVVVDNRGSGHRGLKFEFAGFYKMGQIEIDDOVEGLQY	720
Db	661	VNNRFGVKYFRNLNTLASLGYVVVVVVDNRGSGHRGLKFEFAGFYKMGQIEIDDOVEGLQY	720
QY	721	LASRYDFIDLDRVGIHGSYGYLSLALMQRSDIFRVAIAGAPVTLWIFYDTGYTERYM	780
Db	721	LASRYDFIDLDRVGIHGSYGYLSLALMQRSDIFRVAIAGAPVTLWIFYDTGYTERYM	780
QY	781	GHPDQNEQGYILGVSAMQAEKFPPEPNRLLHLLHLYQENLGSRIAALKVI	840
Db	781	GHPDQNEQGYILGVSAMQAEKFPPEPNRLLHLLHLYQENLGSRIAALKVI	840
QY	841	LOIYPQERHSIRVPESGEHYELHLLHLYQENLGSRIAALKVI	882
Db	841	LOIYPQERHSIRVPESGEHYELHLLHLYQENLGSRIAALKVI	882
RESULT 2			
US-10-070-464-1			
; Sequence 1, Application US/10070464			
; Patent No. 6881564			
; GENERAL INFORMATION:			
; APPLICANT: ABBOTT, Catherine Anne			
; APPLICANT: GORRELL, Mark Douglas			
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES			
; FILE REFERENCE: GH-007			
; CURRENT APPLICATION NUMBER: US/10/070,464			
; CURRENT FILING DATE: 2002-03-07			
; PRIOR APPLICATION NUMBER: PCT/AU00/01085			
; PRIOR FILING DATE: 2000-09-11			
; PRIOR APPLICATION NUMBER: AU PQ5709			
; PRIOR FILING DATE: 2000-02-18			
; PRIOR APPLICATION NUMBER: AU PQ2762			
; PRIOR FILING DATE: 1999-09-10			
; NUMBER OF SEQ ID NOS: 8			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 1			
; LENGTH: 882			
; TYPE: PRT			
; ORGANISM: Homo Sapiens			
US-10-070-464-1			
Query Match 100.0%; Score 4700; DB 2; Length 882;			
Best Local Similarity 100.0%; Pred No. 0;			
Matches 882; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	MAAAMETEQLGVEIFETADCEENIESODPKLEPFYVERYSWSQKLLADTRKYHGYYM	60
Db	1	MAAAMETEQLGVEIFETADCEENIESODPKLEPFYVERYSWSQKLLADTRKYHGYYM	60
QY	61	AKAPHDFMFKVRNDPDPGPHSDRIYYIAMSGENRENTLFI SEIPKTIINRAAVMLSKPPL	120

RESULT 3
US-09-976-674-19
; Sequence 19, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19

LENGTH: 658
TYPE: PRT
ORGANISM: Homo sapiens
US-09-976-674-19

Query Match 74.6%; Score 3504; DB 2; Length 658;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 655; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 1 MAAAMETEOLGVEIFETADCEENIESQDRPKLEPFYVERYSWSQKLLADTRKHYGMM 60
Db 1 MAAAMETEOLGVEIFETADCEENIESQDRPKLEPFYVERYSWSQKLLADTRKHYGMM 60
Qy 61 AKAPHDFMVKNDPDGPHSDRIYYVLAMSGENRENTLFYSEIPKTNINRAAVLMSWKPLL 120
Db 61 AKAPHDFMVKNDPDGPHSDRIYYVLAMSGENRENTLFYSEIPKTNINRAAVLMSWKPLL 120
Qy 121 DLFOATLDYGMYSREBEELLRRKRIGTVGIASYDYGSGTFLFOAGSGIYHVKGQPG 180
Db 121 DLFOATLDYGMYSREBEELLRRKRIGTVGIASYDYGSGTFLFOAGSGIYHVKGQPG 180
Qy 181 FTQOPLRPNLVETSCPNIRMDPKLCPADPDWIAFIHSDNIWISNIVTREERRLLTYVHNL 240
Db 181 FTQOPLRPNLVETSCPNIRMDPKLCPADPDWIAFIHSDNIWISNIVTREERRLLTYVHNL 240
Qy 241 ANNEEDARSAGVATFVLOEEDFDRYSYWCPCAKETTPSGGKILRLIYEENDESEVEI 300
Db 241 ANNEEDARSAGVATFVLOEEDFDRYSYWCPCAKETTPSGGKILRLIYEENDESEVEI 300
Qy 301 TSPMLETRADSFYRKPTGTANPKVTFKMSIMIDAEGRIIDVIDKELIQPFELFEGVE 360
Db 301 TSPMLETRADSFYRKPTGTANPKVTFKMSIMIDAEGRIIDVIDKELIQPFELFEGVE 360
Qy 361 YIARAGWTPEGKYAWSILLDRSQRLQIVLISPELFIPVEDDVNERQRLIESVPDSVTPL 420
Db 361 YIARAGWTPEGKYAWSILLDRSQRLQIVLISPELFIPVEDDVNERQRLIESVPDSVTPL 420
Qy 421 IYEEETDWINIHDFHVPQSHHEEIEFIPASECKTGFRHLKYKITSILKESKYRSSG 480
Db 421 IYEEETDWINIHDFHVPQSHHEEIEFIPASECKTGFRHLKYKITSILKESKYRSSG 480
Qy 481 GLPAPSDFCPIKEEIAITSGEWEVLGRHGSNIQVDEVRRLVYFEGTKDSPLEHLLYVVS 540
Db 481 GLPAPSDFCPIKEEIAITSGEWEVLGRHGSNIQVDEVRRLVYFEGTKDSPLEHLLYVVS 540
Qy 541 YNPNGEVTRLTDRGYSHSCCISQHCDFIFISKYSNQKNPHCVSLYKLSPPDDPTCKTKEF 600
Db 541 YNPNGEVTRLTDRGYSHSCCISQHCDFIFISKYSNQKNPHCVSLYKLSPPDDPTCKTKEF 600
Qy 601 WATILDSAGPLPDYTPPEIFSPESITGFTLYGMLYKPHDLQPGKKYPTVLFYIGG 655
Db 601 WATILDSAGPLPDYTPPEIFSPESITGFTLYGMLYKPHDLQPGKKYPTVLFYIGG 655
```

RESULT 4
US-09-976-674-11

; Sequence 11, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976, 674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 661

TYPE: PRT
ORGANISM: Homo sapiens
US-09-976-674-11

Query Match 74.6%; Score 3504; DB 2; Length 661;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 655; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 1 MAAAMETEOLGVEIFETADCEENIESQDRPKLEPFYVERYSWSQKLLADTRKHYGMM 60
Db 1 MAAAMETEOLGVEIFETADCEENIESQDRPKLEPFYVERYSWSQKLLADTRKHYGMM 60
Qy 61 AKAPHDFMVKNDPDGPHSDRIYYVLAMSGENRENTLFYSEIPKTNINRAAVLMSWKPLL 120
Db 61 AKAPHDFMVKNDPDGPHSDRIYYVLAMSGENRENTLFYSEIPKTNINRAAVLMSWKPLL 120
Qy 121 DLFOATLDYGMYSREBEELLRRKRIGTVGIASYDYGSGTFLFOAGSGIYHVKGQPG 180
Db 121 DLFOATLDYGMYSREBEELLRRKRIGTVGIASYDYGSGTFLFOAGSGIYHVKGQPG 180
Qy 181 FTQOPLRPNLVETSCPNIRMDPKLCPADPDWIAFIHSDNIWISNIVTREERRLLTYVHNL 240
Db 181 FTQOPLRPNLVETSCPNIRMDPKLCPADPDWIAFIHSDNIWISNIVTREERRLLTYVHNL 240
Qy 241 ANNEEDARSAGVATFVLOEEDFDRYSYWCPCAKETTPSGGKILRLIYEENDESEVEI 300
Db 241 ANNEEDARSAGVATFVLOEEDFDRYSYWCPCAKETTPSGGKILRLIYEENDESEVEI 300
Qy 301 TSPMLETRADSFYRKPTGTANPKVTFKMSIMIDAEGRIIDVIDKELIQPFELFEGVE 360
Db 301 TSPMLETRADSFYRKPTGTANPKVTFKMSIMIDAEGRIIDVIDKELIQPFELFEGVE 360
Qy 361 YIARAGWTPEGKYAWSILLDRSQRLQIVLISPELFIPVEDDVNERQRLIESVPDSVTPL 420
Db 361 YIARAGWTPEGKYAWSILLDRSQRLQIVLISPELFIPVEDDVNERQRLIESVPDSVTPL 420
Qy 421 IYEEETDWINIHDFHVPQSHHEEIEFIPASECKTGFRHLKYKITSILKESKYRSSG 480
Db 421 IYEEETDWINIHDFHVPQSHHEEIEFIPASECKTGFRHLKYKITSILKESKYRSSG 480
Qy 481 GLPAPSDFCPIKEEIAITSGEWEVLGRHGSNIQVDEVRRLVYFEGTKDSPLEHLLYVVS 540
Db 481 GLPAPSDFCPIKEEIAITSGEWEVLGRHGSNIQVDEVRRLVYFEGTKDSPLEHLLYVVS 540
Qy 541 YNPNGEVTRLTDRGYSHSCCISQHCDFIFISKYSNQKNPHCVSLYKLSPPDDPTCKTKEF 600
Db 541 YNPNGEVTRLTDRGYSHSCCISQHCDFIFISKYSNQKNPHCVSLYKLSPPDDPTCKTKEF 600
Qy 601 WATILDSAGPLPDYTPPEIFSPESITGFTLYGMLYKPHDLQPGKKYPTVLFYIGG 655
Db 601 WATILDSAGPLPDYTPPEIFSPESITGFTLYGMLYKPHDLQPGKKYPTVLFYIGG 655
```

RESULT 5

US-09-976-674-7
Sequence 7, Application US/09976674
Patent No. 6844180

GENERAL INFORMATION:
APPLICANT: Qi, Steve
APPLICANT: Akinsanya, Karen
APPLICANT: Riviere, Pierre
APPLICANT: Junien, Jean-Louis
TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
FILE REFERENCE: 70669
CURRENT APPLICATION NUMBER: US/09/976, 674
CURRENT FILING DATE: 2001-10-12
PRIOR APPLICATION NUMBER: US 60/240,117
PRIOR FILING DATE: 2000-10-12
NUMBER OF SEQ ID NOS: 61
SOFTWARE: PatentIn version 3.1
SEQ ID NO 7
LENGTH: 690
TYPE: PRT

US-09-976-674-7
; ORGANISM: Homo sapiens
Query Match
Best Local Similarity 74.6%; Score 3504; DB 2; Length 690;
Matches 655; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAAAMETQLGVETPETAADCEENIESQDRPKLPFFVYVERYSWSQLKLADTRKYHGYMM 60
DB 1 MAAAMETQLGVETPETAADCEENIESQDRPKLPFFVYVERYSWSQLKLADTRKYHGYMM 60
QY 61 AKAPHDFMFKRNDPDGPHSDRIYYLAMSGENRENTLFYSEIPKTIINRAAVLMLSWKPL 120
DB 61 AKAPHDFMFKRNDPDGPHSDRIYYLAMSGENRENTLFYSEIPKTIINRAAVLMLSWKPL 120
QY 121 DLFOATLDYGMYSREELLRERKIGTVGIASVDYHQSCTFLFOAGSGIYHVKDGGPG 180
DB 121 DLFOATLDYGMYSREELLRERKIGTVGIASVDYHQSCTFLFOAGSGIYHVKDGGPG 180
QY 181 FTQOPLRPNLVETSCPNIRMDPKLCPADPDWIAFIHNSDIWISNIVTREERLTYVHNL 240
DB 181 FTQOPLRPNLVETSCPNIRMDPKLCPADPDWIAFIHNSDIWISNIVTREERLTYVHNL 240
QY 241 ANMEEDARSAGVATFVLQEEFDRYSGYWCPCKAETTPSGGKILRIIYEENDESEVEI 300
DB 241 ANMEEDARSAGVATFVLQEEFDRYSGYWCPCKAETTPSGGKILRIIYEENDESEVEI 300
QY 301 TSPMLETRRADSFYRYPKTTGTANPKVTFMSEIMIDAGRIIDVIDKELIQPFILFEGVE 360
DB 301 TSPMLETRRADSFYRYPKTTGTANPKVTFMSEIMIDAGRIIDVIDKELIQPFILFEGVE 360
QY 361 YIARAGWTPGKYAWSILLDRSOTRLQIVLISPELFIPVEDDDYMERQRIESVPDSVTPL 420
DB 361 YIARAGWTPGKYAWSILLDRSOTRLQIVLISPELFIPVEDDDYMERQRIESVPDSVTPL 420
QY 421 IIVEETTDIWINIHDIHFVFPQSHEEIEFIPASECTGPRHLKYKITSILKESKYRSG 480
DB 421 IIVEETTDIWINIHDIHFVFPQSHEEIEFIPASECTGPRHLKYKITSILKESKYRSG 480
QY 481 GLPAPSDFKCPKEIAITSGEWEVLGRHGSNIQVDEVRRLVYFEGTKDSPLEHLLYVVS 540
DB 481 GLPAPSDFKCPKEIAITSGEWEVLGRHGSNIQVDEVRRLVYFEGTKDSPLEHLLYVVS 540
QY 541 YVNPGEVRLTDRGYSHSCCISQHCDFPISKYSNQKNPHCVSLYKLSPPEDDPTCKTKEF 600
DB 541 YVNPGEVRLTDRGYSHSCCISQHCDFPISKYSNQKNPHCVSLYKLSPPEDDPTCKTKEF 600
QY 601 WATILDSAGPLPDYTPPEIFSFESTTGTLYGMLYKPHDLQPGKKYPTVLFIYGG 655
DB 601 WATILDSAGPLPDYTPPEIFSFESTTGTLYGMLYKPHDLQPGKKYPTVLFIYGG 655

RESULT 6
US-09-976-674-21
; Sequence 21, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 21
; LENGTH: 613
; TYPE: PRT
; ORGANISM: Homo sapiens

US-09-976-674-21
Query Match
Best Local Similarity 68.9%; Score 3236; DB 2; Length 613;
Matches 607; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAAAMETQLGVETPETAADCEENIESQDRPKLPFFVYVERYSWSQLKLADTRKYHGYMM 60
DB 1 MAAAMETQLGVETPETAADCEENIESQDRPKLPFFVYVERYSWSQLKLADTRKYHGYMM 60
QY 61 AKAPHDFMFKRNDPDGPHSDRIYYLAMSGENRENTLFYSEIPKTIINRAAVLMLSWKPL 120
DB 61 AKAPHDFMFKRNDPDGPHSDRIYYLAMSGENRENTLFYSEIPKTIINRAAVLMLSWKPL 120
QY 121 DLFOATLDYGMYSREELLRERKIGTVGIASVDYHQSCTFLFOAGSGIYHVKDGGPG 180
DB 121 DLFOATLDYGMYSREELLRERKIGTVGIASVDYHQSCTFLFOAGSGIYHVKDGGPG 180
QY 181 FTQOPLRPNLVETSCPNIRMDPKLCPADPDWIAFIHNSDIWISNIVTREERLTYVHNL 240
DB 181 FTQOPLRPNLVETSCPNIRMDPKLCPADPDWIAFIHNSDIWISNIVTREERLTYVHNL 240
QY 241 ANMEEDARSAGVATFVLQEEFDRYSGYWCPCKAETTPSGGKILRIIYEENDESEVEI 300
DB 241 ANMEEDARSAGVATFVLQEEFDRYSGYWCPCKAETTPSGGKILRIIYEENDESEVEI 300
QY 301 TSPMLETRRADSFYRYPKTTGTANPKVTFMSEIMIDAGRIIDVIDKELIQPFILFEGVE 360
DB 301 TSPMLETRRADSFYRYPKTTGTANPKVTFMSEIMIDAGRIIDVIDKELIQPFILFEGVE 360
QY 361 YIARAGWTPGKYAWSILLDRSOTRLQIVLISPELFIPVEDDDYMERQRIESVPDSVTPL 420
DB 361 YIARAGWTPGKYAWSILLDRSOTRLQIVLISPELFIPVEDDDYMERQRIESVPDSVTPL 420
QY 421 IIVEETTDIWINIHDIHFVFPQSHEEIEFIPASECTGPRHLKYKITSILKESKYRSG 480
DB 421 IIVEETTDIWINIHDIHFVFPQSHEEIEFIPASECTGPRHLKYKITSILKESKYRSG 480
QY 481 GLPAPSDFKCPKEIAITSGEWEVLGRHGSNIQVDEVRRLVYFEGTKDSPLEHLLYVVS 540
DB 481 GLPAPSDFKCPKEIAITSGEWEVLGRHGSNIQVDEVRRLVYFEGTKDSPLEHLLYVVS 540
QY 541 YVNPGEVRLTDRGYSHSCCISQHCDFPISKYSNQKNPHCVSLYKLSPPEDDPTCKTKEF 600
DB 541 YVNPGEVRLTDRGYSHSCCISQHCDFPISKYSNQKNPHCVSLYKLSPPEDDPTCKTKEF 600
QY 601 WATILDS 607
DB 601 WATILDS 607
RESULT 7
US-09-976-674-3
; Sequence 3, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 3
; LENGTH: 863
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-3


```
QY 498 ITSGEWVLGRGSHNIQVDEVRRLVYFEGTKDSPLEHLLVYVYVNPGEVTLTRDGYSH 557
Db 181 ITSGEWVLGRGSHNIQVDEVRRLVYFEGTKDSPLEHLLVYVYVNPGEVTLTRDGYSH 240
QY 558 SCCISOHCDPFIKYSNQKPHCVSLYKLSPPEDDPTCKTKFWMATILDSAGPLPDYTPP 617
Db 241 SCCISOHCDPFIKYSNQKPHCVSLYKLSPPEDDPTCKTKFWMATILDSAGPLPDYTPP 300
QY 618 EIPFSESTGTGLTGLMYLKHDLQPGKKYPTVLFYIGGQVOLVNNRFGKVKYFRLNTLA 677
Db 301 EIPFSESTGTGLTGLMYLKHDLQPGKKYPTVLFYIGGQ----- 340
QY 678 SLGYVVVVIDNRGSHRGKLFEGAFKYMGOIIDDQVEGLQYLASRYDFIDLDRVGTHG 737
Db 341 ----- 340
QY 738 WSYGGYLSLMMALMORSIDIFRAIAGAPVTLWIFDYTGTYERYMGHPDQNEQYILGSVAM 797
Db 341 -----VAIAGAPVTLWIFDYTGTYERYMGHPDQNEQYILGSVAM 380
QY 798 QAEKFPSPENRLLHGLFDBNVHFAHTSLLSFLVRAGKPYDLQIYPOERHSIRVPESG 857
Db 381 QAEKFPSPENRLLHGLFDBNVHFAHTSLLSFLVRAGKPYDLQIYPOERHSIRVPESG 440
QY 858 EHYELHLLHYLOENLGRSRIAAKVJ 882
Db 441 EHYELHLLHYLOENLGRSRIAAKVJ 465
```

```
RESULT 13
US-09-976-674-29
; Sequence 29, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; PRIOR FILING DATE: 2001-10-12
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 29
; LENGTH: 832
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-29
```

```
Query Match 51.2%; Score 2406; DB 2; Length 832;
Best Local Similarity 59.0%; Pred. No. 3.2e-236;
Matches 434; Conservative 122; Mismatches 177; Indels 2; Gaps 2;

QY 35 FVYRYSWSQLKLLADTRKYHYMMAKAPHDFMVKRNDPDGPHSDRIYYLAMSGENRE 94
Db 53 FQVQKHSWDGLRSIIHGSRKYSGLIVNKAHPDFQVQKTDGSPHSHRLYILGMPYGSRE 112
QY 95 NTLFYSEIPKTNRAAVALMSWKPLDLFQATLDYGMYSREELLERKRKRIGTVGIASYD 154
Db 113 NSLLYSEIPKPKVRKEALLLSWKQMLDHFQATPHHGVSREELLERKRKRIGTVGIASYD 172
QY 155 YHOGSGTFLFQAGSGIYHVKGDPGQGTQQPLRPNLVETSCNIRMDPKLCPADPDWIAF 214
Db 173 FHSRSGTLFQASNSLFCRDRGKNGFMVSPKPLKTCQSPRMDPKICPADPAFTSF 232
QY 215 THSNIDWISNTVTRERRLTVVHNLNMEEDARSAGVATFVLQEFDFRYSYWCPCXAE 274
Db 233 INNSDLWANIETGEERLTFCHQGLSNVLDPKSAGVATFVLQEFDFRYSYWCPCXAE 292
```

```
QY 275 TTPSGG-KILRLLYEENDESEVEIIVHTSPMLERADSPRYPKTGTANPKVTFKMSIEM 333
Db 293 WEGSGGLTKLRLLYEENDESEVEIIVHVPSPALBEKRTDSYRPTGSKNPKIALKLAEPQ 352
QY 334 IDAERIIDVDIKELIQBFELIFEGVEYIARAGWTPEGKYAMSIILDRSQTRLOIYLISP 393
Db 353 TDSQKIVSTQEKELVQFSSLPFKVYIARAGWTRDGTANAWFLDRPOOMLQVLVLLPP 412
QY 394 ELIFIVEDVMERQRLIESVPSDVTPLIIYBEETDIWINHIDIHVPVPOSH-BEEIEFTF 452
Db 413 ALUFISTNEQRQASARAVRNQPVYVVEEVNWINVHIDIFYPPQSEGEDELCEFLR 472
QY 453 ASECKTGRHLYKITSILSKYSKSSGGLPAPSDFKCPIKEBIAITSGSEWELGRHGSN 512
Db 473 ANECKTGCHLYKYTAVLKQSGYDMSBFPSPGDEDFKCPIKEBIAITSGSEWELARHGSK 532
QY 513 IOVDEVRRLVYFEGTKDSPLEHLLVYVYVNPGEVTLTRDGYSHSCCSIQHCDFFISKY 572
Db 533 IWNSETKLTVYFQGTQDTPLEHLLVYVYVYEAAGEIVRLTTPGFSHSCSMSONDFMESHY 592
QY 573 SNOKNPHCVSLYKLSPPEDDPTCKTKFWMATILDSAGPLPDYTPPPIFSPSTTGLTYG 632
Db 593 SSVSTPPCVHVYKLSGPDPLHKKQPRFASMMEAASCPDPYVVPBIFPHFTRSDVRLYG 652
QY 633 MLYKPHDLQPGKKYPTVLFYIGGQVOLVNNRFGKVKYFRLNTLASLGYVVVVIDNRGSC 692
Db 653 MLYKPHDLQPGKKYPTVLFYIGGQVOLVNNRFGKVKYFRLNTLASLGYVVVVIDNRGSC 712
QY 693 HRGLFEGAFKYMGOIIDDQVEGLQYLASRYDFIDLDRVGTHGSHSGYLSLMMALMQR 752
Db 713 QRGLRFEKALNKGQVEIDQVEGLQFVAKYGFIDLSRVAIHGWSYGGFLSLMGLIHK 772
QY 753 SDIFRAIAGAPVTL 767
Db 773 PQVFAQPLAYPPRL 787
```

```
RESULT 14
US-09-976-674-31
; Sequence 31, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 31
; LENGTH: 832
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-31
```

```
Query Match 51.2%; Score 2406; DB 2; Length 832;
Best Local Similarity 59.0%; Pred. No. 3.2e-236;
Matches 434; Conservative 122; Mismatches 177; Indels 2; Gaps 2;

QY 35 FVYRYSWSQLKLLADTRKYHYMMAKAPHDFMVKRNDPDGPHSDRIYYLAMSGENRE 94
Db 53 FQVQKHSWDGLRSIIHGSRKYSGLIVNKAHPDFQVQKTDGSPHSHRLYILGMPYGSRE 112
QY 95 NTLFYSEIPKTNRAAVALMSWKPLDLFQATLDYGMYSREELLERKRKRIGTVGIASYD 154
Db 113 NSLLYSEIPKPKVRKEALLLSWKQMLDHFQATPHHGVSREELLERKRKRIGTVGIASYD 172
QY 155 YHOGSGTFLFQAGSGIYHVKGDPGQGTQQPLRPNLVETSCNIRMDPKLCPADPDWIAF 214
```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 14, 2006, 13:01:45 ; Search time 172 Seconds
(without alignments)
2142.590 Million cell updates/sec

Title: US-10-825-632-1
Perfect score: 4700
Sequence: 1 MAAAMETEQLGVETFTADC.....HLHLQLQENLGRKIAALKVI 882

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA Main.*
1: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pep.*
3: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pep.*
4: /cgn2_6/prodata/1/pubpaa/US10A_PUBCOMB.pep.*
5: /cgn2_6/prodata/1/pubpaa/US10B_PUBCOMB.pep.*
6: /cgn2_6/prodata/1/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	4700	100.0	882	3	US-09-976-674-1
2	4700	100.0	882	4	US-10-054-776-2
3	4700	100.0	882	4	US-10-170-789-38
4	4700	100.0	882	4	US-10-311-035-9
5	4700	100.0	882	4	US-10-072-012-622
6	4700	100.0	882	4	US-10-415-122-6
7	4700	100.0	882	4	US-10-825-632-1
8	4700	100.0	882	5	US-10-982-512-1
9	4528.5	96.4	883	4	US-10-072-012-621
10	3504	74.6	658	3	US-09-976-674-19
11	3504	74.6	658	5	US-10-982-512-19
12	3504	74.6	661	3	US-09-976-674-11
13	3504	74.6	661	5	US-10-982-512-11
14	3504	74.6	690	3	US-09-976-674-7
15	3504	74.6	690	5	US-10-982-512-7
16	3236	68.9	613	3	US-09-976-674-21
17	3236	68.9	613	5	US-10-982-512-21
18	2870	61.1	863	3	US-09-976-674-3
19	2870	61.1	863	5	US-10-982-512-3
20	2870	61.1	863	5	US-10-982-512-23
21	2870	61.1	892	3	US-09-976-674-27
22	2870	61.1	892	5	US-10-982-512-27
23	2870	61.1	892	5	US-10-982-512-27
24	2870	61.1	892	5	US-10-433-757-12
25	2870	61.1	892	5	US-10-415-122-2
26	2863	60.9	969	4	US-10-1072-012-224
27	2862	60.9	863	4	US-10-072-012-224

ALIGNMENTS

RESULT 1

US-09-976-674-1
; Sequence 1, Application US/09976674
; Patent No. US20020115843A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-1

Query Match	100.0%;	Score 4700;	DB 3;	Length 882;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 882;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MAAAMETEQLGVETFTADCENIESQDRPKLEPPFVVERYSQKLKLLADTRKYHYGYM	60	
DB	1	MAAAMETEQLGVETFTADCENIESQDRPKLEPPFVVERYSQKLKLLADTRKYHYGYM	60	
QY	61	AKAHPDMFKVRNDPDPGPHSDRIYYLLAMSGENRENTLFYSEIPKTNRAAVLMLSWKPLL	120	
DB	61	AKAHPDMFKVRNDPDPGPHSDRIYYLLAMSGENRENTLFYSEIPKTNRAAVLMLSWKPLL	120	
QY	121	DLFQATLDYGMYSREELLRRKRI GTVGIASYDHYHQSGTFLFOAGSGIYHVKGQGGQ	180	
DB	121	DLFQATLDYGMYSREELLRRKRI GTVGIASYDHYHQSGTFLFOAGSGIYHVKGQGGQ	180	
QY	181	FTQQLPRLNVLVETSCPNIRMDPKLCPADPDWIAFIHNSNDIWSNIVTREERLTYYVHNEL	240	
DB	181	FTQQLPRLNVLVETSCPNIRMDPKLCPADPDWIAFIHNSNDIWSNIVTREERLTYYVHNEL	240	
QY	241	ANMEDARSAGATVFLQEEFDYSGYWCPCAKETTPSGGKILRIIYENDESEVEI IHV	300	
DB	241	ANMEDARSAGATVFLQEEFDYSGYWCPCAKETTPSGGKILRIIYENDESEVEI IHV	300	
QY	301	TSPMLTERRADSPRYPKTGTANPKVTFMSEIMDAEGRIIDVIDKELIOPFEILFEGVE	360	
DB	301	TSPMLTERRADSPRYPKTGTANPKVTFMSEIMDAEGRIIDVIDKELIOPFEILFEGVE	360	

```
QY 361 YIARAGWTPGKYAWSILLDRSQTRLQIVLISPELFIPVEDDVMERQRLIESVPDSVTP 420
Db 361 YIARAGWTPGKYAWSILLDRSQTRLQIVLISPELFIPVEDDVMERQRLIESVPDSVTP 420
QY 421 IYEEETDIWINIHDIHFVFPQSHEEIEIFASECKTGRHLKYKITSILKESKYRSG 480
Db 421 IYEEETDIWINIHDIHFVFPQSHEEIEIFASECKTGRHLKYKITSILKESKYRSG 480
QY 481 GLPAPSDFKCPIKEBIAITSGEWEVLGRHGSNIQVDEVRLVYFEGTKDSPLEHLYVVS 540
Db 481 GLPAPSDFKCPIKEBIAITSGEWEVLGRHGSNIQVDEVRLVYFEGTKDSPLEHLYVVS 540
QY 541 YNPGEVTRLDRGYSHSCCISQCHDFFIISKYSNQKNPHCVSLYKLSPEDDPTCKTKEF 600
Db 541 YNPGEVTRLDRGYSHSCCISQCHDFFIISKYSNQKNPHCVSLYKLSPEDDPTCKTKEF 600
QY 601 WATILDSAGPLPDYTPPEIFSFESTTGTLYGMLYKPHDLQPGKYPTVLFYGGPQVOL 660
Db 601 WATILDSAGPLPDYTPPEIFSFESTTGTLYGMLYKPHDLQPGKYPTVLFYGGPQVOL 660
QY 661 VNNRFGVKYFRNLNTLASLGYVVVVVVDNRGSGCHRGKLFEGAFKYKMGQIEIDDOVEGLQY 720
Db 661 VNNRFGVKYFRNLNTLASLGYVVVVVVDNRGSGCHRGKLFEGAFKYKMGQIEIDDOVEGLQY 720
QY 721 LASRYDFDLDVRGHHGWSYGGYLSLALMQRSDIFRVAIAGAPVTLWIFYDTGYTERYM 780
Db 721 LASRYDFDLDVRGHHGWSYGGYLSLALMQRSDIFRVAIAGAPVTLWIFYDTGYTERYM 780
QY 781 GHPDNEQGYLYGVSVAQAEKFPSEPNRLLLLHGFLDENVHFAHTSILLSPFLVRAGKPYD 840
Db 781 GHPDNEQGYLYGVSVAQAEKFPSEPNRLLLLHGFLDENVHFAHTSILLSPFLVRAGKPYD 840
QY 841 LQIYQERHSIRVPESGHEHLLHYLQENLGSRIAALKVI 882
Db 841 LQIYQERHSIRVPESGHEHLLHYLQENLGSRIAALKVI 882
```

RESULT 2

```
US-10-054-776-2
; Sequence 2, Application US/10054776
; Publication No. US2003016581A1
; GENERAL INFORMATION:
; APPLICANT: Mark Robert Edbrooke
; APPLICANT: Alan Peter Lewis
; TITLE OF INVENTION: NOVEL PROTEIN
; FILE REFERENCE: QG1042US
; CURRENT APPLICATION NUMBER: US/10/054,776
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-054-776-2
```

```
Query Match
Best Local Similarity 100.0%; Score 4700; DB 4; Length 882;
Matches 882; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAANMETBOLGVEIFETADCEENIESODRPKLEBPFVVERYSOLKLLADTRKYHYMM 60
Db 1 MAANMETBOLGVEIFETADCEENIESODRPKLEBPFVVERYSOLKLLADTRKYHYMM 60
QY 61 AKAPHDPMFKVRNDPDGHSRIYLLAWSGENRENTLFYSEIPKTNRAAVLMSWKILL 120
Db 61 AKAPHDPMFKVRNDPDGHSRIYLLAWSGENRENTLFYSEIPKTNRAAVLMSWKILL 120
QY 121 DLFOATLDYGYMSREBELLRKRKRGITGVGIASVDYHGGSGTFLFQAGSGIYHVKGQGG 180
Db 121 DLFOATLDYGYMSREBELLRKRKRGITGVGIASVDYHGGSGTFLFQAGSGIYHVKGQGG 180
```

```
QY 181 FTQOPLRPNLVETSPNIRMDPKLCPADPDWIAPIHSNDIWIINIVTREERRLTYVHNEL 240
Db 181 FTQOPLRPNLVETSPNIRMDPKLCPADPDWIAPIHSNDIWIINIVTREERRLTYVHNEL 240
QY 241 ANNEEDARSAGVATVFLQEEEDRYSGYVWCPKAEKTPSGGKIIRILYEENDESEVERI 300
Db 241 ANNEEDARSAGVATVFLQEEEDRYSGYVWCPKAEKTPSGGKIIRILYEENDESEVERI 300
QY 301 TSPMLETRADSFRIYPTKGTANTPKVTPKMSIMIDAEGRIIDVIDKELIQPFILPEGVE 360
Db 301 TSPMLETRADSFRIYPTKGTANTPKVTPKMSIMIDAEGRIIDVIDKELIQPFILPEGVE 360
QY 361 YIARAGWTPGKYAWSILLDRSQTRLQIVLISPELFIPVEDDVMERQRLIESVPDSVTP 420
Db 361 YIARAGWTPGKYAWSILLDRSQTRLQIVLISPELFIPVEDDVMERQRLIESVPDSVTP 420
QY 421 IYEEETDIWINIHDIHFVFPQSHEEIEIFASECKTGRHLKYKITSILKESKYRSG 480
Db 421 IYEEETDIWINIHDIHFVFPQSHEEIEIFASECKTGRHLKYKITSILKESKYRSG 480
QY 481 GLPAPSDFKCPIKEBIAITSGEWEVLGRHGSNIQVDEVRLVYFEGTKDSPLEHLYVVS 540
Db 481 GLPAPSDFKCPIKEBIAITSGEWEVLGRHGSNIQVDEVRLVYFEGTKDSPLEHLYVVS 540
QY 541 YNPGEVTRLDRGYSHSCCISQCHDFFIISKYSNQKNPHCVSLYKLSPEDDPTCKTKEF 600
Db 541 YNPGEVTRLDRGYSHSCCISQCHDFFIISKYSNQKNPHCVSLYKLSPEDDPTCKTKEF 600
QY 601 WATILDSAGPLPDYTPPEIFSFESTTGTLYGMLYKPHDLQPGKYPTVLFYGGPQVOL 660
Db 601 WATILDSAGPLPDYTPPEIFSFESTTGTLYGMLYKPHDLQPGKYPTVLFYGGPQVOL 660
QY 661 VNNRFGVKYFRNLNTLASLGYVVVVVVDNRGSGCHRGKLFEGAFKYKMGQIEIDDOVEGLQY 720
Db 661 VNNRFGVKYFRNLNTLASLGYVVVVVVDNRGSGCHRGKLFEGAFKYKMGQIEIDDOVEGLQY 720
QY 721 LASRYDFDLDVRGHHGWSYGGYLSLALMQRSDIFRVAIAGAPVTLWIFYDTGYTERYM 780
Db 721 LASRYDFDLDVRGHHGWSYGGYLSLALMQRSDIFRVAIAGAPVTLWIFYDTGYTERYM 780
QY 781 GHPDNEQGYLYGVSVAQAEKFPSEPNRLLLLHGFLDENVHFAHTSILLSPFLVRAGKPYD 840
Db 781 GHPDNEQGYLYGVSVAQAEKFPSEPNRLLLLHGFLDENVHFAHTSILLSPFLVRAGKPYD 840
QY 841 LQIYQERHSIRVPESGHEHLLHYLQENLGSRIAALKVI 882
Db 841 LQIYQERHSIRVPESGHEHLLHYLQENLGSRIAALKVI 882
```

RESULT 3

```
US-10-170-789-38
; Sequence 38, Application US/10170789
; Publication No. US20030180930A1
; GENERAL INFORMATION:
; APPLICANT: Rachel E. Meyers
; APPLICANT: Olandt, Peter J.
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Williamson, Mark
; APPLICANT: Weich, Nadine
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASE, PHOSPHATASE,
; FILE REFERENCE: 10448-191001
; CURRENT APPLICATION NUMBER: US/10/170,789
; CURRENT FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: US 09/797,039
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: PCI/US01/06525
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/186,061
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 09/882,166
; PRIOR FILING DATE: 2001-06-15
```

```

; PRIOR APPLICATION NUMBER: PCT/US01/19269
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/212,078
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/934,406
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: PCT/US01/26052
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/226,740
; PRIOR FILING DATE: 2000-08-21
; PRIOR APPLICATION NUMBER: US 09/861,801
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: PCT/US01/16549
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: US 60/205,508
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 09/801,267
; PRIOR FILING DATE: 2001-03-06
; PRIOR APPLICATION NUMBER: PCT/US01/07138
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: US 60/187,454
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: US 09/829,671
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: PCT/US01/40483
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: US 60/197,508
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: US 09/961,721
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: PCT/US01/29904
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: US 60/235,023
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US 10/045,367
; PRIOR FILING DATE: 2001-11-07
; PRIOR APPLICATION NUMBER: US 60/246,561
; PRIOR FILING DATE: 2000-11-07
; PRIOR APPLICATION NUMBER: US 09/801,275
; PRIOR FILING DATE: 2001-03-06
; PRIOR APPLICATION NUMBER: PCT/US01/07074
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: US 60/187,420
; PRIOR FILING DATE: 2000-03-07
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-170-789-38

Query Match      100.0%; Score 4700; DB 4; Length 882;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 882; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1  MAAAMETEQLGVEIFETADCEENIESQDRPKLEPPFYVERYSWSOLKLLADTKYHYNM 60
Db      1  MAAAMETEQLGVEIFETADCEENIESQDRPKLEPPFYVERYSWSOLKLLADTKYHYNM 60

Qy      61  AKAPHDFMFVKRNDPDGPHSDRIYYLAMSGENRENTLFYSEIPKTTINRAAVMLSNKPL 120
Db      61  AKAPHDFMFVKRNDPDGPHSDRIYYLAMSGENRENTLFYSEIPKTTINRAAVMLSNKPL 120

Qy      121  DLFOATLDYGMYSREELLREKRIGTVGIASDYHOGSGTFLFOAGSGIYHVKGDPQG 180
Db      121  DLFOATLDYGMYSREELLREKRIGTVGIASDYHOGSGTFLFOAGSGIYHVKGDPQG 180

Qy      181  FTQOPLRPNLVETSCFNIRMDPKLPADPDWIAFTHSNDIWSINIVTREERRLTYYHNE 240
Db      181  FTQOPLRPNLVETSCFNIRMDPKLPADPDWIAFTHSNDIWSINIVTREERRLTYYHNE 240

Qy      241  ANNEEDARSAGVATFVLQEEPRDYSYWWCPKAEITTSGGKILRLIYEENDESEVEIHV 300

```

```

Db      241  ANNEEDARSAGVATFVLQEEPRDYSYWWCPKAEITTSGGKILRLIYEENDESEVEIHV 300
Qy      301  TSPMLETRRADSFYRYPKGTGTANPKVTFKQSEIMIDAEGRIIDVIDKELIQPFELFPGVE 360
Db      301  TSPMLETRRADSFYRYPKGTGTANPKVTFKQSEIMIDAEGRIIDVIDKELIQPFELFPGVE 360
Qy      361  YIARAGWTPEGKYAWSILLDRSQTRLQIIVLSPELFPVEDDVNERQLIESVPDSVTPL 420
Db      361  YIARAGWTPEGKYAWSILLDRSQTRLQIIVLSPELFPVEDDVNERQLIESVPDSVTPL 420
Qy      421  IYVEETTDIWINHDIHFVPOSHEEIEFIPASECTGFRHLKYIISILKESKYKSSG 480
Db      421  IYVEETTDIWINHDIHFVPOSHEEIEFIPASECTGFRHLKYIISILKESKYKSSG 480
Qy      481  GLPAPSDFKCPIKEEIAITSGEWEVLGRHGSNIQVDEVRLVYFEGTKDSPLEHLLYVS 540
Db      481  GLPAPSDFKCPIKEEIAITSGEWEVLGRHGSNIQVDEVRLVYFEGTKDSPLEHLLYVS 540
Qy      541  YNPNGEVTRLDRGYSHSCCISQHCDFPISKYSNQKNPHCVSLYKLSPEDDPTCKTKEF 600
Db      541  YNPNGEVTRLDRGYSHSCCISQHCDFPISKYSNQKNPHCVSLYKLSPEDDPTCKTKEF 600
Qy      601  WATILDSAGPLPDYTPPEIFSPESSTGFTLYGMLYKPHDLQPGKKYPTVLFYGGPOVL 660
Db      601  WATILDSAGPLPDYTPPEIFSPESSTGFTLYGMLYKPHDLQPGKKYPTVLFYGGPOVL 660
Qy      661  VNNRFGVKYFRNLTLASLGYYVVVVDNRRGSHRGKPEGAFKYMGOIEIDDOVEGLQY 720
Db      661  VNNRFGVKYFRNLTLASLGYYVVVVDNRRGSHRGKPEGAFKYMGOIEIDDOVEGLQY 720
Qy      721  LASRYDFIDLDRVGHGWSYGYLSLMALMORSIDFRVAIAGAPVTLMIFDTGYTRYM 780
Db      721  LASRYDFIDLDRVGHGWSYGYLSLMALMORSIDFRVAIAGAPVTLMIFDTGYTRYM 780
Qy      781  GHPDQNEQGYLGSVAMQAEKPEPSEPNRLILHLGFLDENVHFAHTSILLSPVLRAGPYD 840
Db      781  GHPDQNEQGYLGSVAMQAEKPEPSEPNRLILHLGFLDENVHFAHTSILLSPVLRAGPYD 840
Qy      841  LQIYPOERHSIRVPESGEHYELHLLHYLQENLGSRIAALKVI 882
Db      841  LQIYPOERHSIRVPESGEHYELHLLHYLQENLGSRIAALKVI 882

RESULT 4
US-10-311-035-9
; Sequence 9, Application US/10311035
; Publication No. US20040023243A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: YUE, Henry
; APPLICANT: ELLIOTT, Vicki
; APPLICANT: GANDHI, Ameena R.
; APPLICANT: LAU, Preeti
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: TRIBOULEY, Catherine M.
; APPLICANT: DELEGEANE, Angelo M.
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: NGUYEN, Damiel B.
; APPLICANT: LEE, Ernestine A.
; APPLICANT: HAPALIA, April
; APPLICANT: KHAN, Farrah A.
; APPLICANT: CHAWLA, Narinder K.
; APPLICANT: LU, Monique G.
; APPLICANT: LU, Dyrung Aina M.
; APPLICANT: ARVIZU, Chandra S.
; APPLICANT: TANG, Y. Tom
; APPLICANT: WALSH, Roderick T.
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: LU, Yan
; APPLICANT: RAMKUNAR, Jayalaximi
; APPLICANT: XU, Yuming
; APPLICANT: REDDY, Roopa

```

APPLICANT: DAS, Depopriya
APPLICANT: KEARNEY, Liam
APPLICANT: KALLICK, Deborah A.
TITLE OF INVENTION: Proteases
FILE REFERENCE: PI-0123 PCT
CURRENT APPLICATION NUMBER: US/10/311.035
CURRENT FILING DATE: 2002-12-10
PRIOR APPLICATION NUMBER: 60/213,336; 60/213,955; 60/215,396; 60/216,821; 60/218,946
PRIOR FILING DATE: 2000-06-16; 2000-06-22; 2000-06-29; 2000-07-07; 2000-07-14
NUMBER OF SEQ ID NOS: 42
SOFTWARE: PERL Program
SEQ ID NO 9
LENGTH: 882
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. US20040023243A1 7160544CD1
US-10-311-035-9

Query Match 100.0%; Score 4700; DB 4; Length 882;
Best Local Similarity 100.0%; Pred No 0;
Matches 882; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAAMETEGVIEIFETADCENTIESQDRPKLEPPYVERYSMSQLKLLADTRKHYGMM 60
Db 1 MAAMETEGVIEIFETADCENTIESQDRPKLEPPYVERYSMSQLKLLADTRKHYGMM 60
Qy 61 AKAPHDFMVKNDPDGPHSDRIYYLWAMSGENRENTLFYSEIPIKTNRAAVLMSKPLL 120
Db 61 AKAPHDFMVKNDPDGPHSDRIYYLWAMSGENRENTLFYSEIPIKTNRAAVLMSKPLL 120
Qy 121 DLFOATLDYGMYSREELLRKRKIGTVGIASYDHQSGTFLFQAGSGIYHVKGPG 180
Db 121 DLFOATLDYGMYSREELLRKRKIGTVGIASYDHQSGTFLFQAGSGIYHVKGPG 180
Qy 181 FTQOPLRPNLVETSCFNIRMDPKCPADPDWIAPIHNSDIWISNIVTREERLTYVHNEL 240
Db 181 FTQOPLRPNLVETSCFNIRMDPKCPADPDWIAPIHNSDIWISNIVTREERLTYVHNEL 240
Qy 241 ANNEEDARSAGVATFVLOEEDFDRSGYMWCPKAEITTPSGGKILRLIYEENDESEVEIHV 300
Db 241 ANNEEDARSAGVATFVLOEEDFDRSGYMWCPKAEITTPSGGKILRLIYEENDESEVEIHV 300
Qy 301 TSPMLETRRADSFYRYPKGTANPKVTFKMSIIMIDASGRIIDVIDKELIOPPEILPEGVE 360
Db 301 TSPMLETRRADSFYRYPKGTANPKVTFKMSIIMIDASGRIIDVIDKELIOPPEILPEGVE 360
Qy 361 YIARAGWTPEGKYANSILLDRSOTELQIVLISPELFIPEVDDVNERQLIESVPDSVTPL 420
Db 361 YIARAGWTPEGKYANSILLDRSOTELQIVLISPELFIPEVDDVNERQLIESVPDSVTPL 420
Qy 421 IYEEETDWINIHDFVFPQSHHEERTIFASECKTGRHLYKITSILKESKYKSSG 480
Db 421 IYEEETDWINIHDFVFPQSHHEERTIFASECKTGRHLYKITSILKESKYKSSG 480
Qy 481 GLPAPSDFKCPIKEEIAITSGEWEVLGRHGSNIQDEVRLIYFEGKDSPLEHHLYVVS 540
Db 481 GLPAPSDFKCPIKEEIAITSGEWEVLGRHGSNIQDEVRLIYFEGKDSPLEHHLYVVS 540
Qy 541 YVNPGEVTRTLDRGSHSCCISQHCDFIFISKYSNQKPHCVSLYKLSPPDDCTCKTKEF 600
Db 541 YVNPGEVTRTLDRGSHSCCISQHCDFIFISKYSNQKPHCVSLYKLSPPDDCTCKTKEF 600
Qy 601 WATILDSAGFLPDYTPPEIFSPFESTTGTGLYGLYKPHDLQPGKKYPTVLIYGGPQVQL 660
Db 601 WATILDSAGFLPDYTPPEIFSPFESTTGTGLYGLYKPHDLQPGKKYPTVLIYGGPQVQL 660
Qy 661 VNNRFGVKYFRNLTLASLGYYVVVIDNRGSHRGLKPEGAFAFKYKMGQIIDDQVEGLQY 720
Db 661 VNNRFGVKYFRNLTLASLGYYVVVIDNRGSHRGLKPEGAFAFKYKMGQIIDDQVEGLQY 720
Qy 721 LASKRYDFIDLDVRGIGHWSYGGYLSLMALMORSIDIFRVAIAGAPVTLWIFDYTGTYERYM 780

Db 721 LASKRYDFIDLDVRGIGHWSYGGYLSLMALMORSIDIFRVAIAGAPVTLWIFDYTGTYERYM 780
Qy 781 GHPDQNEQYVLYGSVAMQAEKFPSEPNRLILLHGFLDENVHFAHTSILLSFVLRACKPYD 840
Db 781 GHPDQNEQYVLYGSVAMQAEKFPSEPNRLILLHGFLDENVHFAHTSILLSFVLRACKPYD 840
Qy 841 LQIYQERHSIRVPESGHEHYELHLLHYLOENLGSRIIAALKVI 882
Db 841 LQIYQERHSIRVPESGHEHYELHLLHYLOENLGSRIIAALKVI 882

RESULT 5

US-10-072-012-622
; Sequence 622, Application US/10072012
; Publication No. US20040033493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberley
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Shinkete, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Bsha
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier, Jr. Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Colman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grosse, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 622
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-072-012-622

CURRENT APPLICATION NUMBER: US/10/415.122
 CURRENT FILING DATE: 2003-08-07
 NUMBER OF SEQ ID NOS: 8
 SOFTWARE: Patent in version 3.1

Query Match 100.0%; Score 4700; DB 4; Length 882;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 882; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MAAMETEOLGVEIIFETADCEENISQDRPKLEPPYVERYSQKLLADTRKYHGYMM 60
 1 MAAMETEOLGVEIIFETADCEENISQDRPKLEPPYVERYSQKLLADTRKYHGYMM 60
 61 AKAPHDPMFVKRNDPDPGPHSDRIYYLAMSNGENRENTLFYSEIPKTIINRAAVMLSWKPLL 120
 61 AKAPHDPMFVKRNDPDPGPHSDRIYYLAMSNGENRENTLFYSEIPKTIINRAAVMLSWKPLL 120
 121 DLFOATLDYGMYSREBELLRERKRICTVGIASDYHQSGLTFQAGSGIYHVKGDPQG 180
 121 DLFOATLDYGMYSREBELLRERKRICTVGIASDYHQSGLTFQAGSGIYHVKGDPQG 180
 181 FTQOPLRPNLVETSCNIRMDPKLCPADPDWIAFHSNDIWI SNIVTREERLTYVHNEL 240
 181 FTQOPLRPNLVETSCNIRMDPKLCPADPDWIAFHSNDIWI SNIVTREERLTYVHNEL 240
 241 ANNEEDARSAGVATFVLOEFDRYSGYMWCPKAEITPSGGKILRIIYEENDESEVEIIVH 300
 241 ANNEEDARSAGVATFVLOEFDRYSGYMWCPKAEITPSGGKILRIIYEENDESEVEIIVH 300
 301 TSPMLETRRADSPRYPKTGTANPKVTFKMSIIMDAEGRIIDVIDKELIOPFEILFEGVE 360
 301 TSPMLETRRADSPRYPKTGTANPKVTFKMSIIMDAEGRIIDVIDKELIOPFEILFEGVE 360
 361 YIARAGWTPGKYAWSILLDRSOTRLOIQLVISPFLIPVEDDVMERQRLIESVPDSVTPL 420
 361 YIARAGWTPGKYAWSILLDRSOTRLOIQLVISPFLIPVEDDVMERQRLIESVPDSVTPL 420
 421 IYEBETTDIWINIHDIHVPFQSHHEEIEFIFASECKTGFRLHYKITSILKESKYKSSG 480
 421 IYEBETTDIWINIHDIHVPFQSHHEEIEFIFASECKTGFRLHYKITSILKESKYKSSG 480
 481 GLPAPSDFKCPIKEEIAITSGEWEVLGRHGSNIQVDEVRRLVYFEGTKDSPLEHLYVVS 540
 481 GLPAPSDFKCPIKEEIAITSGEWEVLGRHGSNIQVDEVRRLVYFEGTKDSPLEHLYVVS 540
 541 YNPGEVTRLDRGYSHSCCIHQCDFFISKYSNQKNPHCVSLYKLSPPEDDPTCKTKEF 600
 541 YNPGEVTRLDRGYSHSCCIHQCDFFISKYSNQKNPHCVSLYKLSPPEDDPTCKTKEF 600
 601 WATILDSAGPLDPTTPEIFSFESTTGTLYGMLYKPHDLQPKKYPTVLFYGGPOVOL 660
 601 WATILDSAGPLDPTTPEIFSFESTTGTLYGMLYKPHDLQPKKYPTVLFYGGPOVOL 660
 661 VNNRFGVKYFRNLNTLASGLYVVVVVINDRGSCHRGKLFEGAFKYMGOIEIDDOVEGLQY 720
 661 VNNRFGVKYFRNLNTLASGLYVVVVVINDRGSCHRGKLFEGAFKYMGOIEIDDOVEGLQY 720
 721 LASRYDFDLDVRGTHGWSYGYLSLMAIMQKSDIFRVAIAGAPVTLMIFYDTGYTERYM 780
 721 LASRYDFDLDVRGTHGWSYGYLSLMAIMQKSDIFRVAIAGAPVTLMIFYDTGYTERYM 780
 781 GHPDQNEQGYLLGSVAMQAEKFPSEPNRLLHGLFDENVHFAHTSILLSFLVRACKPYD 840
 781 GHPDQNEQGYLLGSVAMQAEKFPSEPNRLLHGLFDENVHFAHTSILLSFLVRACKPYD 840
 841 LQIYPOERHSIRVPESGEHYELHLLHYLOENLGSRIAALKVI 882
 841 LQIYPOERHSIRVPESGEHYELHLLHYLOENLGSRIAALKVI 882

RESULT 6

US-10-415-122-6
 ; Sequence 6, Application US/10415122
 ; Publication No. US20040053369A1
 ; GENERAL INFORMATION:
 ; APPLICANT: THE UNIVERSITY OF SYDNEY
 ; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES
 ; FILE REFERENCE: FP15127

1 MAAMETEOLGVEIIFETADCEENISQDRPKLEPPYVERYSQKLLADTRKYHGYMM 60
 1 MAAMETEOLGVEIIFETADCEENISQDRPKLEPPYVERYSQKLLADTRKYHGYMM 60
 61 AKAPHDPMFVKRNDPDPGPHSDRIYYLAMSNGENRENTLFYSEIPKTIINRAAVMLSWKPLL 120
 61 AKAPHDPMFVKRNDPDPGPHSDRIYYLAMSNGENRENTLFYSEIPKTIINRAAVMLSWKPLL 120
 121 DLFOATLDYGMYSREBELLRERKRICTVGIASDYHQSGLTFQAGSGIYHVKGDPQG 180
 121 DLFOATLDYGMYSREBELLRERKRICTVGIASDYHQSGLTFQAGSGIYHVKGDPQG 180
 181 FTQOPLRPNLVETSCNIRMDPKLCPADPDWIAFHSNDIWI SNIVTREERLTYVHNEL 240
 181 FTQOPLRPNLVETSCNIRMDPKLCPADPDWIAFHSNDIWI SNIVTREERLTYVHNEL 240
 241 ANNEEDARSAGVATFVLOEFDRYSGYMWCPKAEITPSGGKILRIIYEENDESEVEIIVH 300
 241 ANNEEDARSAGVATFVLOEFDRYSGYMWCPKAEITPSGGKILRIIYEENDESEVEIIVH 300
 301 TSPMLETRRADSPRYPKTGTANPKVTFKMSIIMDAEGRIIDVIDKELIOPFEILFEGVE 360
 301 TSPMLETRRADSPRYPKTGTANPKVTFKMSIIMDAEGRIIDVIDKELIOPFEILFEGVE 360
 361 YIARAGWTPGKYAWSILLDRSOTRLOIQLVISPFLIPVEDDVMERQRLIESVPDSVTPL 420
 361 YIARAGWTPGKYAWSILLDRSOTRLOIQLVISPFLIPVEDDVMERQRLIESVPDSVTPL 420
 421 IYEBETTDIWINIHDIHVPFQSHHEEIEFIFASECKTGFRLHYKITSILKESKYKSSG 480
 421 IYEBETTDIWINIHDIHVPFQSHHEEIEFIFASECKTGFRLHYKITSILKESKYKSSG 480
 481 GLPAPSDFKCPIKEEIAITSGEWEVLGRHGSNIQVDEVRRLVYFEGTKDSPLEHLYVVS 540
 481 GLPAPSDFKCPIKEEIAITSGEWEVLGRHGSNIQVDEVRRLVYFEGTKDSPLEHLYVVS 540
 541 YNPGEVTRLDRGYSHSCCIHQCDFFISKYSNQKNPHCVSLYKLSPPEDDPTCKTKEF 600
 541 YNPGEVTRLDRGYSHSCCIHQCDFFISKYSNQKNPHCVSLYKLSPPEDDPTCKTKEF 600
 601 WATILDSAGPLDPTTPEIFSFESTTGTLYGMLYKPHDLQPKKYPTVLFYGGPOVOL 660
 601 WATILDSAGPLDPTTPEIFSFESTTGTLYGMLYKPHDLQPKKYPTVLFYGGPOVOL 660
 661 VNNRFGVKYFRNLNTLASGLYVVVVVINDRGSCHRGKLFEGAFKYMGOIEIDDOVEGLQY 720
 661 VNNRFGVKYFRNLNTLASGLYVVVVVINDRGSCHRGKLFEGAFKYMGOIEIDDOVEGLQY 720
 721 LASRYDFDLDVRGTHGWSYGYLSLMAIMQKSDIFRVAIAGAPVTLMIFYDTGYTERYM 780
 721 LASRYDFDLDVRGTHGWSYGYLSLMAIMQKSDIFRVAIAGAPVTLMIFYDTGYTERYM 780
 781 GHPDQNEQGYLLGSVAMQAEKFPSEPNRLLHGLFDENVHFAHTSILLSFLVRACKPYD 840
 781 GHPDQNEQGYLLGSVAMQAEKFPSEPNRLLHGLFDENVHFAHTSILLSFLVRACKPYD 840
 841 LQIYPOERHSIRVPESGEHYELHLLHYLOENLGSRIAALKVI 882
 841 LQIYPOERHSIRVPESGEHYELHLLHYLOENLGSRIAALKVI 882

```

RESULT 7
US-10-825-632-1
; Sequence 1, Application US/10825632
; Publication No. US20040191826A1
; GENERAL INFORMATION:
; APPLICANT: ABBOTT, Catherine Anne
; APPLICANT: GORRELL, Mark Douglas
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES
; FILE REFERENCE: FCSB-100-Div. 1
; CURRENT APPLICATION NUMBER: US/10/825,632
; CURRENT FILING DATE: 2004-04-15
; PRIOR APPLICATION NUMBER: US 10/070,464
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: PCT/AU00/01085
; PRIOR FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: AU P05709
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: AU P02762
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-825-632-1

Query Match      100.0%; Score 4700; DB 4; Length 882;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 882; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAAMETEQLGVEIFETADCEENIESQDRPKLEPPFVVERYSWSQKLLADTRKXHYGM 60
DB 1 MAAAMETEQLGVEIFETADCEENIESQDRPKLEPPFVVERYSWSQKLLADTRKXHYGM 60
QY 61 AKAPHPDMFKVRNDPDGPHSDRIYVLAMSGENRENTLFYSEIPKTIINRAAVLMSWKPLL 120
DB 61 AKAPHPDMFKVRNDPDGPHSDRIYVLAMSGENRENTLFYSEIPKTIINRAAVLMSWKPLL 120
QY 121 DLFOATLDYGMYSREBEELLREKRIGTVGIASVDYHOGSGTFLFOAGSGIYHVKGDPQG 180
DB 121 DLFOATLDYGMYSREBEELLREKRIGTVGIASVDYHOGSGTFLFOAGSGIYHVKGDPQG 180
QY 181 FTQOPLRPNLVETSCPNIRMDPKLCPADPDWIAFIHSDNIWISNIVTREERLTYVHNEL 240
DB 181 FTQOPLRPNLVETSCPNIRMDPKLCPADPDWIAFIHSDNIWISNIVTREERLTYVHNEL 240
QY 241 ANNEEDARSAGVATFVLQEEFDYSGYWCPCAEITTPSGGKILRIIYEENDESEVEIHW 300
DB 241 ANNEEDARSAGVATFVLQEEFDYSGYWCPCAEITTPSGGKILRIIYEENDESEVEIHW 300
QY 301 TSPMLETRADSFYPKTGTAHPKVTFFKMSIIMDAEGRIIDVIDKELIQPFILFEGVE 360
DB 301 TSPMLETRADSFYPKTGTAHPKVTFFKMSIIMDAEGRIIDVIDKELIQPFILFEGVE 360

RESULT 8
US-10-982-512-1
; Sequence 1, Application US/10982512
; Publication No. US20050059081A1
; GENERAL INFORMATION:
; APPLICANT: Oi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Juuinen, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/10/982,512
; CURRENT FILING DATE: 2004-11-05
; PRIOR APPLICATION NUMBER: US/09/976,674
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-982-512-1

Query Match      100.0%; Score 4700; DB 5; Length 882;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 882; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAAMETEQLGVEIFETADCEENIESQDRPKLEPPFVVERYSWSQKLLADTRKXHYGM 60
DB 1 MAAAMETEQLGVEIFETADCEENIESQDRPKLEPPFVVERYSWSQKLLADTRKXHYGM 60
QY 61 AKAPHPDMFKVRNDPDGPHSDRIYVLAMSGENRENTLFYSEIPKTIINRAAVLMSWKPLL 120
DB 61 AKAPHPDMFKVRNDPDGPHSDRIYVLAMSGENRENTLFYSEIPKTIINRAAVLMSWKPLL 120
QY 121 DLFOATLDYGMYSREBEELLREKRIGTVGIASVDYHOGSGTFLFOAGSGIYHVKGDPQG 180
DB 121 DLFOATLDYGMYSREBEELLREKRIGTVGIASVDYHOGSGTFLFOAGSGIYHVKGDPQG 180
QY 181 FTQOPLRPNLVETSCPNIRMDPKLCPADPDWIAFIHSDNIWISNIVTREERLTYVHNEL 240
DB 181 FTQOPLRPNLVETSCPNIRMDPKLCPADPDWIAFIHSDNIWISNIVTREERLTYVHNEL 240
QY 241 ANNEEDARSAGVATFVLQEEFDYSGYWCPCAEITTPSGGKILRIIYEENDESEVEIHW 300
DB 241 ANNEEDARSAGVATFVLQEEFDYSGYWCPCAEITTPSGGKILRIIYEENDESEVEIHW 300
QY 301 TSPMLETRADSFYPKTGTAHPKVTFFKMSIIMDAEGRIIDVIDKELIQPFILFEGVE 360
DB 301 TSPMLETRADSFYPKTGTAHPKVTFFKMSIIMDAEGRIIDVIDKELIQPFILFEGVE 360

```

QY	361	YIARAGWTPEGKYAWSIILLDRSOTRLQIVLISPELFIPEVDDVMEORLIESVPDSVTP	420
Db	361	YIARAGWTPEGKYAWSIILLDRSOTRLQIVLISPELFIPEVDDVMEORLIESVPDSVTP	420
QY	421	LIYEETDIWINIHDI PHVFPQSHEEIEFI FASECTGPRHLKYKITSILKESKYRSSG	480
Db	421	LIYEETDIWINIHDI PHVFPQSHEEIEFI FASECTGPRHLKYKITSILKESKYRSSG	480
QY	481	GUPAPSDPKPIKEBIATITSGEWEVLGRHGSNIQVDEVRRLVYFEGTKDSPLEHLYVVS	540
Db	481	GUPAPSDPKPIKEBIATITSGEWEVLGRHGSNIQVDEVRRLVYFEGTKDSPLEHLYVVS	540
QY	541	YVNPGEVTRLTDRGYSHSCCISQHCDFPISKYSNQKNPHCVSLYKLSPPEDDPTCKTKEF	600
Db	541	YVNPGEVTRLTDRGYSHSCCISQHCDFPISKYSNQKNPHCVSLYKLSPPEDDPTCKTKEF	600
QY	601	WATILDSAGPLPDYTPPEIFSPFESTTGFTLYGMLYKPHDLQPGKKYPTVLFIYGGPQVL	660
Db	601	WATILDSAGPLPDYTPPEIFSPFESTTGFTLYGMLYKPHDLQPGKKYPTVLFIYGGPQVL	660
QY	661	VNNRPGVKYFRLNTLASIGYVVVVIDNRGSGCHRLKFEPAFKYKMGQIBIDQVEGLQY	720
Db	661	VNNRPGVKYFRLNTLASIGYVVVVIDNRGSGCHRLKFEPAFKYKMGQIBIDQVEGLQY	720
QY	721	LASRYDFIDLDVRGIGHWSYGYLSLMALMORSIDIFRVAIAGAPVTILWIFYDTGYTERYM	780
Db	721	LASRYDFIDLDVRGIGHWSYGYLSLMALMORSIDIFRVAIAGAPVTILWIFYDTGYTERYM	780
QY	781	GHPDQNEQGYLLGSVAMQAEKFPSEPNRLLLHGHFLDENVHPAHTSILLSLFLVRACKPYD	840
Db	781	GHPDQNEQGYLLGSVAMQAEKFPSEPNRLLLHGHFLDENVHPAHTSILLSLFLVRACKPYD	840
QY	841	LQIYPOERHSIRVPESGEHYELHLHYLOENLGRIAALKVI	882
Db	841	LQIYPOERHSIRVPESGEHYELHLHYLOENLGRIAALKVI	882
RESULT 9			
US-10-072-012-621			
; Sequence 621, Application US/10072012			
; Publication No. US20040033493A1			
; GENERAL INFORMATION:			
; APPLICANT: Tchernev, Velizar			
; APPLICANT: Spytek, Kimberly			
; APPLICANT: Zernhusen, Bryan			
; APPLICANT: Patturajan, Meera			
; APPLICANT: Shimkets, Richard			
; APPLICANT: Li, Li			
; APPLICANT: Gangolli, Esha			
; APPLICANT: Padigar, Muralidhara			
; APPLICANT: Anderson, David W.			
; APPLICANT: Rastelli, Luca			
; APPLICANT: Miller, Charles E.			
; APPLICANT: Gerlach, Valerie			
; APPLICANT: Taupier Jr, Raymond J.			
; APPLICANT: Gusev, Vladimir Y.			
; APPLICANT: Colman, Steven D.			
; APPLICANT: Wolenc, Adam R.			
; APPLICANT: Pena, Carol E. A.			
; APPLICANT: Furtak, Katarzyna			
; APPLICANT: Grosse, William M.			
; APPLICANT: Alsobrook II, John P.			
; APPLICANT: Lepley, Denise M.			
; APPLICANT: Rieger, Daniel K.			
; APPLICANT: Burgess, Catherine E.			
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same			
; FILE REFERENCE: 21402-258			
; CURRENT APPLICATION NUMBER: US/10/072,012			
; PRIOR FILING DATE: 2002-01-31			
; PRIOR APPLICATION NUMBER: 60/265,102			
; PRIOR FILING DATE: 2001-01-30			
; PRIOR APPLICATION NUMBER: 60/265,514			

PRIOR FILING DATE: 2001-01-31

PRIOR APPLICATION NUMBER: 60/265,517

PRIOR FILING DATE: 2001-01-31

PRIOR APPLICATION NUMBER: 60/265,412

PRIOR FILING DATE: 2001-01-31

PRIOR APPLICATION NUMBER: 60/265,395

PRIOR FILING DATE: 2001-01-31

PRIOR APPLICATION NUMBER: 60/266,406

PRIOR FILING DATE: 2001-02-02

PRIOR APPLICATION NUMBER: 60/266,767

PRIOR FILING DATE: 2001-02-05

PRIOR APPLICATION NUMBER: 60/267,057

PRIOR FILING DATE: 2001-02-07

PRIOR APPLICATION NUMBER: 60/266,975

PRIOR FILING DATE: 2001-02-07

PRIOR APPLICATION NUMBER: 60/267,459

PRIOR FILING DATE: 2001-02-08

PRIOR APPLICATION NUMBER: 60/267,459

Remaining Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 1391

SOFTWARE: Patent in Ver. 2.1

SEQ ID NO 621

LENGTH: 883

TYPE: PRT

ORGANISM: Mus musculus

US-10-072-012-621

Query Match 96.4%; Score 4528.5; DB 4; Length 883;

Best Local Similarity 95.7%; Pred. No. 0;

Matches 845; Conservative 18; Mismatches 19; Indels 1; Gaps 1;

QY	1	MAAAMETEOLGVEIPEITACDEB-NIESODRPKLEPFYVERYSWOLKLLADTRKYGYM	59
Db	1	MAAAMETEOLGVEIPEITACDEB-NIESODRPKLEPFYVERYSWOLKLLADTRKYGYM	60
QY	60	MAKAPHDFMFKVENDPDGPHSDRIYYLAWSGENRENTLIFYSEIPIKTINRAAVLMSKPL	119
Db	61	MAKAPHDFMFKVENDPDGPHSDRIYYLAWSGENRENTLIFYSEIPIKTINRAAVLMSKPL	120
QY	120	LDLFQATLDYGMYSREELLRERKRIGTVGIAADYHOGSGTFLFQAGSGIYHVKGDPQ	179
Db	121	LDLFQATLDYGMYSREELLRERKRIGTVGIAADYHOGSGTFLFQAGSGIYHVKGDPQ	180
QY	180	GFTQQPLRPNLVETSCPNIRMDPKLCPADPDWIAFHNSNDIWNIVTREERRLYVHNE	239
Db	181	GFTQQPLRPNLVETSCPNIRMDPKLCPADPDWIAFHNSNDIWNIVTREERRLYVHNE	240
QY	240	LANNEEDARSAGVATTVLQSEEDRYSGYWCPCAKETTPSGGKILRLIYEENDESEVEI	299
Db	241	LANNEEDARSAGVATTVLQSEEDRYSGYWCPCAKETTPSGGKILRLIYEENDESEVEI	300
QY	300	VTSPLMLETTRADSFYRYPKTGTANPKVTFKMSIIMIDAEGRIIDVIDKELIQPFILPEGV	359
Db	301	VTSPLMLETTRADSFYRYPKTGTANPKVTFKMSIIMIDAEGRIIDVIDKELIQPFILPEGV	360
QY	360	EYIARAGWTPGKYAWSIILLDRSOTRLQIVLISPELFIPEVDDVMEORLIESVPDSVTP	419
Db	361	EYIARAGWTPGKYAWSIILLDRSOTRLQIVLISPELFIPEVDDVMEORLIESVPDSVTP	420
QY	420	LIYEETDIWINIHDI PHVFPQSHEEIEFI FASECTGPRHLKYKITSILKESKYRSS	479
Db	421	LIYEETDIWINIHDI PHVFPQSHEEIEFI FASECTGPRHLKYKITSILKESKYRSS	480
QY	480	GGLPAPSDPKPIKEBIATITSGEWEVLGRHGSNIQVDEVRRLVYFEGTKDSPLEHLYVVS	539
Db	481	GGLPAPSDPKPIKEBIATITSGEWEVLGRHGSNIQVDEVRRLVYFEGTKDSPLEHLYVVS	540
QY	540	SYVNPGEVTRLTDRGYSHSCCISQHCDFPISKYSNQKNPHCVSLYKLSPPEDDPTCKTKE	599
Db	541	SYVNPGEVTRLTDRGYSHSCCISQHCDFPISKYSNQKNPHCVSLYKLSPPEDDPTCKTKE	600
QY	600	FWATILDSAGPLPDYTPPEIFSPFESTTGFTLYGMLYKPHDLQPGKKYPTVLFIYGGPQV	659
Db	601	FWATILDSAGPLPDYTPPEIFSPFESTTGFTLYGMLYKPHDLQPGKKYPTVLFIYGGPQV	660

Qy 660 LVNNRFGKGVKFRNLNTLASLGYYVVVVDNRGSRHGLKFGAPKYKMGQIEIDDOVEGLQ 719
Db 661 LVNNRFGKGVKFRNLNTLASLGYYVVVVDNRGSRHGLKFGAPKYKMGQIEIDDOVEGLQ 720
Qy 720 YLASRYDFTDLDRVGIHGSYGGYLSLMLMQRSDIPRVAIAGAPVTLWIFYDTGYTERY 779
Db 721 YLASQYDFTDLDRVGIHGSYGGYLSLMLMQRSDIPRVAIAGAPVTLWIFYDTGYTERY 780
Qy 780 MGHPPQNEQYGLGSVMAQAKFPSEPNRLLHLLHGLFDENVHFAHTSILLSLFLVRAGKPY 839
Db 781 MGHPPQNEQYGLGSVMAQAKFPSEPNRLLHLLHGLFDENVHFAHTSILLSLFLVRAGKPY 840
Qy 840 DLQIYQERHSIRVPESGEHVELHLLHVLQENLGSRTAALKVI 882
Db 841 DLQIYQERHSIRVPESGEHVELHLLHVLQENLGSRTAALKVI 883

RESULT 10

US-09-976-674-19
; Sequence 19, Application US/09976674
; Publication No. US20020115843A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 19
; LENGTH: 658
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-19

Query Match 74.6%; Score 3504; DB 3; Length 658;
Best Local Similarity 100.0%; Pred. No. 7.3e-308;
Matches 655; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MAAAMETEOLGVEIFETADCEENIESQDRPKLEPPYVERYSWSQKLLADTRKHYGMM 60
Db 1 MAAAMETEOLGVEIFETADCEENIESQDRPKLEPPYVERYSWSQKLLADTRKHYGMM 60
Qy 61 AKAPHDFMVKRNDPDGPHSDRIYYLAMSGENRENTLYFSEIPIKTINRAAVLMSWKPLL 120
Db 61 AKAPHDFMVKRNDPDGPHSDRIYYLAMSGENRENTLYFSEIPIKTINRAAVLMSWKPLL 120
Qy 121 DLFQATLDYGMYSREELLRRKRIGTVGIASDYHOGSGTFLFQAGSGIYHVKGQPG 180
Db 121 DLFQATLDYGMYSREELLRRKRIGTVGIASDYHOGSGTFLFQAGSGIYHVKGQPG 180
Qy 181 FTQOPLRPNLVETSCPNIRMDPKLCPADPWIAFIHSNDIWSNIVTREERLTYVHNEL 240
Db 181 FTQOPLRPNLVETSCPNIRMDPKLCPADPWIAFIHSNDIWSNIVTREERLTYVHNEL 240
Qy 241 ANNEEDARSAGVATFVLOEEDRYSGYWCPCAKETTPSGGKILRIIYEENDESEVEIHHV 300
Db 241 ANNEEDARSAGVATFVLOEEDRYSGYWCPCAKETTPSGGKILRIIYEENDESEVEIHHV 300
Qy 301 TSPMLTRRADSFYPKTGTPANPKVTFKMSIIMDAEGRIIDVIDKELIQPFILFEGVE 360
Db 301 TSPMLTRRADSFYPKTGTPANPKVTFKMSIIMDAEGRIIDVIDKELIQPFILFEGVE 360
Qy 361 YIARAGWTPEGKYAWSILLDRSQTRLQIVLISPELFIPEVDVNMERQRLIESVPDSVTPL 420
Db 361 YIARAGWTPEGKYAWSILLDRSQTRLQIVLISPELFIPEVDVNMERQRLIESVPDSVTPL 420

Qy 421 LIYBETTDIWINIHDIHFVPOSHERRIERIPASECKTGRHLYKITSLIKESKYKSSG 480
Db 421 LIYBETTDIWINIHDIHFVPOSHERRIERIPASECKTGRHLYKITSLIKESKYKSSG 480
Qy 481 GLPAPSDPKPIKEBIATISGEWELGRHGSNIQVDEVRLVYFEGTKOSPLEHHLYVVS 540
Db 481 GLPAPSDPKPIKEBIATISGEWELGRHGSNIQVDEVRLVYFEGTKOSPLEHHLYVVS 540
Qy 541 YNPGGEVTRLDRGYSHSCCISQCHDFFISKYSNQKNPHCVSLYKJLSSPDDPTCKTKGF 600
Db 541 YNPGGEVTRLDRGYSHSCCISQCHDFFISKYSNQKNPHCVSLYKJLSSPDDPTCKTKGF 600
Qy 601 WATILDSAGPLPDYTPPEIFSFESTTGTLYGMLYKPHDLQPGKKYPTVLFYGG 655
Db 601 WATILDSAGPLPDYTPPEIFSFESTTGTLYGMLYKPHDLQPGKKYPTVLFYGG 655

RESULT 11

US-10-982-512-19
; Sequence 19, Application US/10982512
; Publication No. US20050059081A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/10/982,512
; CURRENT FILING DATE: 2004-11-05
; PRIOR APPLICATION NUMBER: US/09/976,674
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 19
; LENGTH: 658
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-982-512-19

Query Match 74.6%; Score 3504; DB 5; Length 658;
Best Local Similarity 100.0%; Pred. No. 7.3e-308;
Matches 655; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MAAAMETEOLGVEIFETADCEENIESQDRPKLEPPYVERYSWSQKLLADTRKHYGMM 60
Db 1 MAAAMETEOLGVEIFETADCEENIESQDRPKLEPPYVERYSWSQKLLADTRKHYGMM 60
Qy 61 AKAPHDFMVKRNDPDGPHSDRIYYLAMSGENRENTLYFSEIPIKTINRAAVLMSWKPLL 120
Db 61 AKAPHDFMVKRNDPDGPHSDRIYYLAMSGENRENTLYFSEIPIKTINRAAVLMSWKPLL 120
Qy 121 DLFQATLDYGMYSREELLRRKRIGTVGIASDYHOGSGTFLFQAGSGIYHVKGQPG 180
Db 121 DLFQATLDYGMYSREELLRRKRIGTVGIASDYHOGSGTFLFQAGSGIYHVKGQPG 180
Qy 181 FTQOPLRPNLVETSCPNIRMDPKLCPADPWIAFIHSNDIWSNIVTREERLTYVHNEL 240
Db 181 FTQOPLRPNLVETSCPNIRMDPKLCPADPWIAFIHSNDIWSNIVTREERLTYVHNEL 240
Qy 241 ANNEEDARSAGVATFVLOEEDRYSGYWCPCAKETTPSGGKILRIIYEENDESEVEIHHV 300
Db 241 ANNEEDARSAGVATFVLOEEDRYSGYWCPCAKETTPSGGKILRIIYEENDESEVEIHHV 300
Qy 301 TSPMLTRRADSFYPKTGTPANPKVTFKMSIIMDAEGRIIDVIDKELIQPFILFEGVE 360
Db 301 TSPMLTRRADSFYPKTGTPANPKVTFKMSIIMDAEGRIIDVIDKELIQPFILFEGVE 360
Qy 361 YIARAGWTPEGKYAWSILLDRSQTRLQIVLISPELFIPEVDVNMERQRLIESVPDSVTPL 420
Db 361 YIARAGWTPEGKYAWSILLDRSQTRLQIVLISPELFIPEVDVNMERQRLIESVPDSVTPL 420

QY 421 IYETTTDWINIHDIHVFPQSHHEIEIFASECKTGRHLKYKITSILKESKYKRS 480
DB 421 IYETTTDWINIHDIHVFPQSHHEIEIFASECKTGRHLKYKITSILKESKYKRS 480
QY 481 GLPAPSDPKCPKEEIAITSGEWEVLGRHGSNIQVDEVRLRVYFEGTKDSPLEHLYVVS 540
DB 481 GLPAPSDPKCPKEEIAITSGEWEVLGRHGSNIQVDEVRLRVYFEGTKDSPLEHLYVVS 540
QY 541 YNPGEVTRLDRGYSHSCCISQHCDFFIISKYSNOKNPHCVSLYKLSPPEDDPTCKTKEF 600
DB 541 YNPGEVTRLDRGYSHSCCISQHCDFFIISKYSNOKNPHCVSLYKLSPPEDDPTCKTKEF 600
QY 601 WATILDSAGPLPDYTPPEIFSFESTTGTLYGMLYKPHDLQPGKKYPTVLFYGG 655
DB 601 WATILDSAGPLPDYTPPEIFSFESTTGTLYGMLYKPHDLQPGKKYPTVLFYGG 655

RESULT 12

US-09-976-674-11
; Sequence 11, Application US/09976674
; Patent No. US20020115843A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; TYPE: PRT
; LENGTH: 661
; ORGANISM: Homo sapiens
US-09-976-674-11

Query Match 74.6%; Score 3504; DB 3; Length 661;
Best Local Similarity 100.0%; Pred. No. 7.3e-308;
Matches 655; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAAMETEQLGVEIFETADCEENIESQDRPKLEPFYVERYSWSQKLLADTRKYHYMM 60
DB 1 MAAAMETEQLGVEIFETADCEENIESQDRPKLEPFYVERYSWSQKLLADTRKYHYMM 60
QY 61 AKAPHDFMFVKRNDPGPHSDRIYVYLAWSGENRENTLFYSEIPKTIINRAAVLMSWKPL 120
DB 61 AKAPHDFMFVKRNDPGPHSDRIYVYLAWSGENRENTLFYSEIPKTIINRAAVLMSWKPL 120
QY 121 DLFOATLDYGMYSREBELLRERKRIGTVGIASDYHQSGTFLFOAGSGIYHVKGDPQG 180
DB 121 DLFOATLDYGMYSREBELLRERKRIGTVGIASDYHQSGTFLFOAGSGIYHVKGDPQG 180
QY 181 FTQOPLRPNLVETSCPNIRMDPKLCPADPDWIAFTHSNDIWSNIVTREERRLTYYHNEL 240
DB 181 FTQOPLRPNLVETSCPNIRMDPKLCPADPDWIAFTHSNDIWSNIVTREERRLTYYHNEL 240
QY 241 ANMEEDARSAGVATFVLQEFDRYSGYMWCPKAETTPSGGKILRLIYEENDESEVEI 300
DB 241 ANMEEDARSAGVATFVLQEFDRYSGYMWCPKAETTPSGGKILRLIYEENDESEVEI 300
QY 301 TSPMLETRRADSPRYPKTGTPANPKVTFKMSIEMDAEGRIIDVIDKELIQPFELLFEGVE 360
DB 301 TSPMLETRRADSPRYPKTGTPANPKVTFKMSIEMDAEGRIIDVIDKELIQPFELLFEGVE 360
QY 361 YIARAGWTPGKYAWSILLDRSQTRQLQIVLISPFLIPVEDDVMERQRLIESVPDSVTPL 420
DB 361 YIARAGWTPGKYAWSILLDRSQTRQLQIVLISPFLIPVEDDVMERQRLIESVPDSVTPL 420

RESULT 13

US-10-982-512-11
; Sequence 11, Application US/10982512
; Publication No. US20050059081A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/10/982,512
; CURRENT FILING DATE: 2004-11-05
; PRIOR APPLICATION NUMBER: US/09/976,674
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 661
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-982-512-11

Query Match 74.6%; Score 3504; DB 5; Length 661;
Best Local Similarity 100.0%; Pred. No. 7.3e-308;
Matches 655; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAAMETEQLGVEIFETADCEENIESQDRPKLEPFYVERYSWSQKLLADTRKYHYMM 60
DB 1 MAAAMETEQLGVEIFETADCEENIESQDRPKLEPFYVERYSWSQKLLADTRKYHYMM 60
QY 61 AKAPHDFMFVKRNDPGPHSDRIYVYLAWSGENRENTLFYSEIPKTIINRAAVLMSWKPL 120
DB 61 AKAPHDFMFVKRNDPGPHSDRIYVYLAWSGENRENTLFYSEIPKTIINRAAVLMSWKPL 120
QY 121 DLFOATLDYGMYSREBELLRERKRIGTVGIASDYHQSGTFLFOAGSGIYHVKGDPQG 180
DB 121 DLFOATLDYGMYSREBELLRERKRIGTVGIASDYHQSGTFLFOAGSGIYHVKGDPQG 180
QY 181 FTQOPLRPNLVETSCPNIRMDPKLCPADPDWIAFTHSNDIWSNIVTREERRLTYYHNEL 240
DB 181 FTQOPLRPNLVETSCPNIRMDPKLCPADPDWIAFTHSNDIWSNIVTREERRLTYYHNEL 240
QY 241 ANMEEDARSAGVATFVLQEFDRYSGYMWCPKAETTPSGGKILRLIYEENDESEVEI 300
DB 241 ANMEEDARSAGVATFVLQEFDRYSGYMWCPKAETTPSGGKILRLIYEENDESEVEI 300
QY 301 TSPMLETRRADSPRYPKTGTPANPKVTFKMSIEMDAEGRIIDVIDKELIQPFELLFEGVE 360
DB 301 TSPMLETRRADSPRYPKTGTPANPKVTFKMSIEMDAEGRIIDVIDKELIQPFELLFEGVE 360
QY 361 YIARAGWTPGKYAWSILLDRSQTRQLQIVLISPFLIPVEDDVMERQRLIESVPDSVTPL 420
DB 361 YIARAGWTPGKYAWSILLDRSQTRQLQIVLISPFLIPVEDDVMERQRLIESVPDSVTPL 420

QY 421 IIVEETTDIWINIHDIHFVFPQSHEEIEFIFASECKTGRHLYKITSILKSKYKRSSG 480
Db 421 IIVEETTDIWINIHDIHFVFPQSHEEIEFIFASECKTGRHLYKITSILKSKYKRSSG 480
QY 481 GLPAPSDKCPKIEKIAITSGEWEVLGRHGSNIQVDEVRLLVYFEGTKDSPLEHLLYVVS 540
Db 481 GLPAPSDKCPKIEKIAITSGEWEVLGRHGSNIQVDEVRLLVYFEGTKDSPLEHLLYVVS 540
QY 541 YNPGEVTRLTDRGYSHSCCIHQCDPFISKYSNQKNPHCVSLYKLSPPEDDPTCKTKEF 600
Db 541 YNPGEVTRLTDRGYSHSCCIHQCDPFISKYSNQKNPHCVSLYKLSPPEDDPTCKTKEF 600
QY 601 WATILDSAGPLPDYTPPEIFSFESTTGTLYGMLYKPHDLQPGKKYPTVLFIYGG 655
Db 601 WATILDSAGPLPDYTPPEIFSFESTTGTLYGMLYKPHDLQPGKKYPTVLFIYGG 655

RESULT 14

US-09-976-674-7
; Sequence 7, Application US/09976674
; Patent No. US20020115843A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 7
; LENGTH: 690
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-7

Query Match 74.6%; Score 3504; DB 3; Length 690;
Best Local Similarity 100.0%; Pred. No. 7.9e-308;
Matches 655; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAAMETEQLGVEIFETADCEENIESQDRPKLEPFYVERYSWSQKLLADTRKYHYWM 60
Db 1 MAAAMETEQLGVEIFETADCEENIESQDRPKLEPFYVERYSWSQKLLADTRKYHYWM 60
QY 61 AKAPHDFMVKRNDPDGPHSDRIYYLAMSGENRENTLIFYSEIPKTNRAAVLMLSKPLL 120
Db 61 AKAPHDFMVKRNDPDGPHSDRIYYLAMSGENRENTLIFYSEIPKTNRAAVLMLSKPLL 120
QY 121 DLFOATLDYGMYSREELLRRKRIGTVGTASDYHOGSGTFLFOAGSGIYHVKGDPQG 180
Db 121 DLFOATLDYGMYSREELLRRKRIGTVGTASDYHOGSGTFLFOAGSGIYHVKGDPQG 180
QY 181 FTQOPLRPNLVETSCPNIRMDPKLCPADPDWIAFIHNSNDIWSNIVTREERLLTYVHNL 240
Db 181 FTQOPLRPNLVETSCPNIRMDPKLCPADPDWIAFIHNSNDIWSNIVTREERLLTYVHNL 240
QY 241 ANNEEDARSAGVATFVLQEEFDYSGYWCPCAKETTPSGGKILRLIYEENDESEVEIHHV 300
Db 241 ANNEEDARSAGVATFVLQEEFDYSGYWCPCAKETTPSGGKILRLIYEENDESEVEIHHV 300
QY 301 TSPMLETRADSFYPKTGTANPKVTFKMSIIMIDABGRIIDVIDKELIQPFELFEGVE 360
Db 301 TSPMLETRADSFYPKTGTANPKVTFKMSIIMIDABGRIIDVIDKELIQPFELFEGVE 360
QY 361 YIARAGWTPGKYAWSILLDRSQRLQIVLISPELFPVDDVMERQRLIESVPDSVTPL 420
Db 361 YIARAGWTPGKYAWSILLDRSQRLQIVLISPELFPVDDVMERQRLIESVPDSVTPL 420

QY 421 IIVEETTDIWINIHDIHFVFPQSHEEIEFIFASECKTGRHLYKITSILKSKYKRSSG 480
Db 421 IIVEETTDIWINIHDIHFVFPQSHEEIEFIFASECKTGRHLYKITSILKSKYKRSSG 480
QY 481 GLPAPSDKCPKIEKIAITSGEWEVLGRHGSNIQVDEVRLLVYFEGTKDSPLEHLLYVVS 540
Db 481 GLPAPSDKCPKIEKIAITSGEWEVLGRHGSNIQVDEVRLLVYFEGTKDSPLEHLLYVVS 540
QY 541 YNPGEVTRLTDRGYSHSCCIHQCDPFISKYSNQKNPHCVSLYKLSPPEDDPTCKTKEF 600
Db 541 YNPGEVTRLTDRGYSHSCCIHQCDPFISKYSNQKNPHCVSLYKLSPPEDDPTCKTKEF 600
QY 601 WATILDSAGPLPDYTPPEIFSFESTTGTLYGMLYKPHDLQPGKKYPTVLFIYGG 655
Db 601 WATILDSAGPLPDYTPPEIFSFESTTGTLYGMLYKPHDLQPGKKYPTVLFIYGG 655

RESULT 15

US-10-982-512-7
; Sequence 7, Application US/10982512
; Publication No. US20050059081A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/10/982,512
; CURRENT FILING DATE: 2004-11-05
; PRIOR APPLICATION NUMBER: US/09/976,674
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 7
; LENGTH: 690
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-982-512-7

Query Match 74.6%; Score 3504; DB 5; Length 690;
Best Local Similarity 100.0%; Pred. No. 7.9e-308;
Matches 655; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAAMETEQLGVEIFETADCEENIESQDRPKLEPFYVERYSWSQKLLADTRKYHYWM 60
Db 1 MAAAMETEQLGVEIFETADCEENIESQDRPKLEPFYVERYSWSQKLLADTRKYHYWM 60
QY 61 AKAPHDFMVKRNDPDGPHSDRIYYLAMSGENRENTLIFYSEIPKTNRAAVLMLSKPLL 120
Db 61 AKAPHDFMVKRNDPDGPHSDRIYYLAMSGENRENTLIFYSEIPKTNRAAVLMLSKPLL 120
QY 121 DLFOATLDYGMYSREELLRRKRIGTVGTASDYHOGSGTFLFOAGSGIYHVKGDPQG 180
Db 121 DLFOATLDYGMYSREELLRRKRIGTVGTASDYHOGSGTFLFOAGSGIYHVKGDPQG 180
QY 181 FTQOPLRPNLVETSCPNIRMDPKLCPADPDWIAFIHNSNDIWSNIVTREERLLTYVHNL 240
Db 181 FTQOPLRPNLVETSCPNIRMDPKLCPADPDWIAFIHNSNDIWSNIVTREERLLTYVHNL 240
QY 241 ANNEEDARSAGVATFVLQEEFDYSGYWCPCAKETTPSGGKILRLIYEENDESEVEIHHV 300
Db 241 ANNEEDARSAGVATFVLQEEFDYSGYWCPCAKETTPSGGKILRLIYEENDESEVEIHHV 300
QY 301 TSPMLETRADSFYPKTGTANPKVTFKMSIIMIDABGRIIDVIDKELIQPFELFEGVE 360
Db 301 TSPMLETRADSFYPKTGTANPKVTFKMSIIMIDABGRIIDVIDKELIQPFELFEGVE 360
QY 361 YIARAGWTPGKYAWSILLDRSQRLQIVLISPELFPVDDVMERQRLIESVPDSVTPL 420
Db 361 YIARAGWTPGKYAWSILLDRSQRLQIVLISPELFPVDDVMERQRLIESVPDSVTPL 420

QY 421 IIVEETTDIWINIHDIHVFPQSHHEEIEFIFASECKTGFRHLYKITSILKESKYKSSG 480
Db 421 IIVEETTDIWINIHDIHVFPQSHHEEIEFIFASECKTGFRHLYKITSILKESKYKSSG 480
QY 481 GLPAPSDFKCPIKEEIAITSGEWELGRHGSNIQDEVRRLVYFEGTKDSPLEHLYVVS 540
Db 481 GLPAPSDFKCPIKEEIAITSGEWELGRHGSNIQDEVRRLVYFEGTKDSPLEHLYVVS 540
QY 541 YVNPGEVTRLTDGYSHSCCISOHCDFFLSKYSNOKNPHCVSLYKLSPPEDDPTCKTKEF 600
Db 541 YVNPGEVTRLTDGYSHSCCISOHCDFFLSKYSNOKNPHCVSLYKLSPPEDDPTCKTKEF 600
QY 601 WATILDSAGPLPDYTPPEIFSPESSTTGFTLYGMLYKPHDLQPGKKYPTVLFYGG 655
Db 601 WATILDSAGPLPDYTPPEIFSPESSTTGFTLYGMLYKPHDLQPGKKYPTVLFYGG 655

Search completed: April 14, 2006, 13:05:34
Job time : 175 secs

SEARCHED APR 18 2006

13:05:34

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 14, 2006, 13:02:54 ; Search time 29 Seconds
(without alignments)
1292.259 Million cell updates/sec

Title: US-10-825-632-1
Perfect score: 4700
Sequence: 1 MAAMWETQLGVETFTADC.....HLHYLQENLGSRIALKVI 882

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 217505 seqs, 42489236 residues

Total number of hits satisfying chosen parameters: 217505

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA New:
1: /STDSS/ptodata/2/pubpaa/US08 NEW PUB.pep.*
2: /STDSS/ptodata/2/pubpaa/US06 NEW PUB.pep.*
3: /STDSS/ptodata/2/pubpaa/US07 NEW PUB.pep.*
4: /STDSS/ptodata/2/pubpaa/ECT_NEW PUB.pep.*
5: /STDSS/ptodata/2/pubpaa/US09 NEW PUB.pep.*
6: /STDSS/ptodata/2/pubpaa/US10 NEW PUB.pep.*
7: /STDSS/ptodata/2/pubpaa/US11 NEW PUB.pep.*
8: /STDSS/ptodata/2/pubpaa/US60 NEW PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4700	100.0	882	7 US-11-151-601-20	Sequence 20, Appl
2	651	13.9	745	7 US-11-079-463-6408	Sequence 6408, Ap
3	577	12.3	109	7 US-11-176-951-10	Sequence 10, Appl
4	551	11.7	627	7 US-11-079-463-7758	Sequence 7758, Ap
5	530	11.3	762	7 US-11-116-939-13	Sequence 13, Appl
6	529	11.3	738	7 US-11-208-288-4	Sequence 4, Appli
7	529	11.3	766	6 US-10-501-035-234	Sequence 234, App
8	529	11.3	766	7 US-11-208-288-2	Sequence 2, Appli
9	525	11.2	766	6 US-10-522-789-2	Sequence 6, Appli
10	517.5	11.0	760	7 US-11-208-288-6	Sequence 55, Appl
11	454	9.7	760	7 US-11-186-284-55	Sequence 11, Appl
12	394	8.4	99	7 US-11-176-951-11	Sequence 1, Appli
13	249.5	5.3	657	7 US-11-179-977-1	Sequence 7504, Ap
14	241.5	5.1	624	7 US-11-079-463-7504	Sequence 9281, Ap
15	200.5	4.3	737	7 US-11-079-463-9281	Sequence 9, Appli
16	145	3.1	102	7 US-11-176-951-7	Sequence 12, Appl
17	139	3.0	115	7 US-11-176-951-9	Sequence 8, Appli
18	139	3.0	115	7 US-11-176-951-12	Sequence 16, Appl
19	133.5	2.8	102	7 US-11-176-951-8	Sequence 982, App
20	125	2.7	24	7 US-11-176-951-16	Sequence 711, Appl
21	119.5	2.5	1019	6 US-10-995-561-982	Sequence 23, Appl
22	119	2.5	1307	6 US-10-995-561-711	Sequence 10550, A
23	118.5	2.5	710	7 US-11-151-601-23	Sequence 1136, Ap
24	118.5	2.5	1155	7 US-11-098-686-10550	
25	118.5	2.5	1243	6 US-10-453-372-1136	

Query Match 100.0%; Score 4700; DB 7; Length 882;
Best Local Similarity 100.0%; Pred. No. 0;

ALIGNMENTS

RESULT 1

US-11-151-601-20
; Sequence 20, Application US/11151601
; Publication No. US20060003413A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Olandt, Peter J.
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Williamson, Mark
; APPLICANT: Welch, Nadine
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASE, PHOSPHATASE,
; AND PROTEASE FAMILY MEMBERS AND USES THEREOF
; FILE REFERENCE: MP100-054P1RCF10MNDIVM
; CURRENT APPLICATION NUMBER: US/11/151,601
; CURRENT FILING DATE: 2005-06-13
; PRIOR APPLICATION NUMBER: US 10/170,789
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: US 09/797,039
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: PCT/US01/06525
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/186,061
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 09/882,166
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: PCT/US01/19269
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/212,078
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/934,406
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: PCT/US01/26052
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/226,740
; PRIOR FILING DATE: 2000-08-21
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-151-601-20

Sequence 1134, Ap
Sequence 1142, Ap
Sequence 1132, Ap
Sequence 18, Appl
Sequence 118, App
Sequence 6314, Ap
Sequence 2, Appli
Sequence 143, App
Sequence 11, Appli
Sequence 3, Appli
Sequence 6, Appli
Sequence 14, Appli
Sequence 16, Appli
Sequence 52, Appli
Sequence 709, App
Sequence 4, Appli
Sequence 6, Appli
Sequence 13, Appli

26 118 2.5 1243 6 US-10-453-372-1134
27 118 2.5 4913 6 US-10-453-372-1142
28 118 2.5 4961 6 US-10-453-372-1132
29 115 2.4 849 6 US-10-909-769-18
30 114.5 2.4 668 6 US-10-454-437-118
31 113.5 2.4 420 7 US-11-079-463-6314
32 113.5 2.4 2483 7 US-11-186-999-2
33 112.5 2.4 1285 7 US-11-206-071-2
34 111 2.4 315 7 US-11-018-868-143
35 110.5 2.4 2458 7 US-11-186-999-11
36 110 2.3 1067 7 US/11/062
37 110 2.3 1092 7 US/11/062
38 110 2.3 2455 7 US-11-186-999-14
39 110 2.3 2455 7 US-11-186-999-16
40 109.5 2.3 2725 7 US-11-113-424-52
41 108.5 2.3 1206 6 US-10-995-561-709
42 108.5 2.3 2256 7 US-11-144-368-4
43 108.5 2.3 2455 7 US-11-186-999-4
44 108.5 2.3 2458 7 US-11-186-999-6
45 108.5 2.3 2458 7 US-11-186-999-13

Db 136 --RQLITERIPN-----NTQWTVSPVGHKLAYVWNNDIYVKIEPNLPSYRITW 183
Qy 236 VHNELANNEEDARAGVATFVLOE--FDRYSGYVWGPCAETTPSGGKILRIIYEENDESE 294
Db 184 TG-----KEDIINYITDWEIEEVSAYSALWNSPNGTF-----LAYAFNOTE 228
Qy 295 VEIIH---VTPSMLETRADSFYKPTGTANPKVTFKMSIIMIDAEGRIIDVIDKELIOP 351
Db 229 VLIIBYSYDESQYKPTKVRVPYKAGAVNPKFV--VNTDSLSSVTNATSQITAP 286
Qy 352 FEILPEGVEYIARAGWTEGKYAWSIILLDRSOTRLQIVLISPELFIPEVDDVMERQRIE 411
Db 287 ASMLI--GDHYLCDVTWA-----TOERISLOWL-----RRIQ 316
Qy 412 SVDPDSVTPLIIEETDIWINIHDFHVPFQSHHEEIEFIFASECKTGPRH-----LY 464
Db 317 NY--SVMDICDYDESSGRW--NC-----LVARQHIEMSTTGWGRPRPSPHFTLDGNGFY 368
Qy 465 KITSILKESKYKRSGLPAPSPDFKCPKEEIAITSEWEVLGRHGSNIQVDEVRRLVVF 524
Db 369 KIIS--NEEGYRHI-----CYFQIDKDCFTITKGTWEVIG-----LEALTSYLYVI 414
Qy 525 EGT--KDSPLEHLYVSVNPGVETRLTDRGYSHSCIS-----OHCDFEISKYSNOKNP 578
Db 415 SNEYKMGPGRNLYKI-----QLSD--YTKVTLSCELNPERCQYYSVSFKEAKY 463
Qy 579 H---C-----VSLYKLSPPEDPTCKTEFWATILDSAGPLDYTPPE-----IFSFESTTGF 628
Db 464 YQRCGGGLPLPLYTLHSSVNDKGLVLED--NSALDKM--LQNVQMPKKLDFIILNETKF 520
Qy 629 TLYGMLYKPHLDQPKKYPVTLFYGGPQVOLVNNRFGVKYFRLN---TLASLGVVVVV 685
Db 521 --WYQMLPPH--FDKSKYPPLLDVYAGPCSQ-----KADTVFLNWAATYLASTENIIVA 572
Qy 686 --IDNRGSHRGLKFEAGFYKMGQIEIDQVQGLQYLASRYDFIDLDRVGIHWSYGGYL 744
Db 573 SFDGRSGYQGDGKIMHAINRRLGTFEVEDQIEAARQF--SKMGFVDNKRITAIWGSYGGYV 631
Qy 745 SLMALMQRSDIFRVAIAGAPVTLWIFDVTGTETERYMG--HPDQNEQGYLGSVAMQAEKF 802
Db 632 TSMVLGSGGVFKCGIAVAPVSRWEYDVSVTERYMGLPTPEDNLHYNSTVMGRAENF 691
Qy 803 PSEPNRLLHGHFDENVHFAHTSILLSFLVRACKPYDLQIYPOERHSIRVPESGHEVEL 862
Db 692 --KQVEYLLIHGTADDNVHFOQSAISKALVDGVDFQAMWYTDDEHGIASSTAQHIIYT 749
Qy 863 HLLHYLOE 870
Db 750 HNSHFQIKQ 757

RESULT 6

US-11-208-288-4
; Sequence 4, Application US/11208288
; Publication No. US20060051366A1
; GENERAL INFORMATION:
; APPLICANT: CHANG, Chiwen
; TITLE OF INVENTION: USE OF SOLUBLE CD26 AS INHIBITOR OF
; FILE REFERENCE: 39533-0001
; CURRENT APPLICATION NUMBER: US/11/208,288
; CURRENT FILING DATE: 2005-08-18
; PRIOR APPLICATION NUMBER: US 60/605,013
; PRIOR FILING DATE: 2004-08-26
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 738
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-208-288-4

Query Match 11.3%; Score 529; DB 7; Length 738;

Best Local Similarity 26.5%; Pred. No. 6.3e-37;
Matches 186; Conservative 109; Mismatches 280; Indels 128; Gaps 32;
Qy 206 PADPDWI-----ARIHNDIWSNIVTREERRLTIVHNELANNEEDARAGVATFV 256
Db 121 PNTQWTVSPVGHKLAYVWNNDIYVKIEPNLPSYRITWGT-----KEDIINYITDWM 174
Qy 257 LORE--FDRYSGYVWGPCAETTPSGGKILRIIYEENDESEVEIIH---VTPSMLETRADS 312
Db 175 YEEVEVSAYSALWNSPNGTF-----LAYAFNOTEVEPLIEYSYDESQYKPTVR 225
Qy 313 FRYKPTGTANPKVTFKMSIIMIDAEGRIIDVIDKELIOPFEILPEGVEYIARAGWTEGK 372
Db 226 VPYKAGAVNPKFV--VNTDSLSSVTNATSQITAPASMLI--GDHYLCDVTWA--- 278
Qy 373 YAWSIILLDRSOTRLQIVLISPELFIPEVDDVMERQRIESVDPDSVTPLIIEETDIWIN 432
Db 279 -----TOERISLOWL-----RRIQNY--SVMDICDYDESSGRW--N 310
Qy 433 IHDIFHVPFQSHHEEIEFIFASECKTGPRH-----LYKITSILKESKYKRSGLPAP 485
Db 311 C-----LVARQHIEMSTTGWGRPRPSPHFTLDGNGFYKIIS--NEEGYRHI----- 356
Qy 486 SDFKCPKEEIAITSEWEVLGRHGSNIQVDEVRRLVVFECT--KDSPLEHLYVSVVNP 544
Db 357 CYFQIDKDCFTITKGTWEVIG-----LEALTSYLYISNEYKMGPGRNLYKIQLIDY 411
Qy 545 GEVTRLTDRGYSHSCIS--OHCDFEISKYSNOKNPH---C-----VSLYKLSPPEDPTCK 596
Db 412 TKVTCL-----SCELNPERCQYYSVSFKEAKYQLRCSGPGPLPLYTLHSSVNDKGLR 464
Qy 597 TKEFWATILDSAGPLDYTPPE---IFSFESTTGTLYGMLYKPHLDQPKKYPVTLFIY 653
Db 465 VLED--NSALDKM--LQNVQMPKKLDFIILNETKF--WYQMLPPH--FDKSKYPPLLDVY 519
Qy 654 GGPQVOLVNNRFGVKYFRLN---TLASLGVVVVV--IDNRGSHRGLKFEAGFYKMGQI 709
Db 520 AGPCSQ-----KADTVFLNWAATYLASTENIIVASFDGRSGYQGDGKIMHAINRRLGTF 573
Qy 710 EIDDQVEGLQYLASRYDFIDLDRVGIHWSYGGYLSMALMQRSDIFRVAIAGAPVTLWI 769
Db 574 EVEDQIEAARQF--SKMGFVDNKRITAIWGSYGGYVTSMLVSGSGGVFKCGIAVAPVSRWE 632
Qy 770 FYDVTGTETERYMG--HPDQNEQGYLGSVAMQAEKFPSEPNRLLHGHFDENVHFAHTSI 827
Db 633 YDVSVTERYMGLPTPEDNLHYNSTVMGRAENF--KQVEYLLIHGTADDNVHFOQSAQ 690
Qy 828 LLSFLVRACKPYDLQIYPOERHSIRVPESGHEVELHLLHYLOE 870
Db 691 ISKALVDGVDFQAMWYTDDEHGIASSTAQHIIYTHMSHFQIKQ 733

RESULT 7

US-10-501-035-234
; Sequence 234, Application US/10501035
; Publication No. US20060046249A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: IDENTIFICATION OF POLYNUCLEOTIDES AND POLYPEPTIDE FOR PREDICTING
; FILE REFERENCE: D0185 PCT
; CURRENT APPLICATION NUMBER: US/10/501,035
; CURRENT FILING DATE: 2004-07-09
; PRIOR APPLICATION NUMBER: US 60/350,061
; PRIOR FILING DATE: 2002-01-18
; NUMBER OF SEQ ID NOS: 795
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 234
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-501-035-234

Query Match 11.3%; Score 529; DB 6; Length 766;
Best Local Similarity 26.5%; Pred. No. 6.7e-37;
Matches 186; Conservative 109; Mismatches 280; Indels 128; Gaps 32;

QY 206 PADPDMW-----AFHSNDIWSNIVTREERRLYVHNLNMEEDARSAGVATFV 256
DB 149 PNTQWTVSPVGHKLAYWNNDIYVKIEPNLPSYRITWTG-----KEDIYNGITDMV 202

QY 257 LQEE-FDRYSGYWCPCAETTPSGGKILRLIYEENDESEVEIHH---VTSPLMETRRADS 312
DB 203 YEEEFVSAYSALWWSNGTF-----LAVAQNDTEVPLIEYSFYSDLSQVPKTVR 253

QY 313 FRYPKGTGTANPKVTFKMSIIMIDAGRIIDVIDKELIQPFIILFEGVEYIARAGWTPGK 372
DB 254 VPYPKAGAVNPTVKFV--VNTDSLSSVTNATSIQITAPASMLI-GDHYLCDVTWA---- 306

QY 373 YAWSILLDRSQRLQIQLVILSPFLFIPVEDDVMERQRLIESVPDSVTPLLIYEETTDIWIN 432
DB 307 -----TQERISLQWL-----RRIQNY--SYMDICDYDESSGRW-N 338

QY 433 IHDIFHVFPQSHHEEIEFIFASECKTGRH-----LYKITSILKESKYKRSSGGLPAP 485
DB 339 C-----LVARQHIEMSTTGWGRFRPSEPHFTLDGNSFYKIIS--NEEGYRHI----- 384

QY 486 SDFKCPKEIAITSGEWELGRHGSNIQVDEVRRLVYFEGT-KDSPLEHHLVYVSVNP 544
DB 385 CYFQIDKKOCTFITKTWEVIG-----IEALTSYLYIISNEYKMGPGGRNLYKIQLIDY 439

QY 545 GEVTRLDRGSHSCCIS--QHCDFFISKYSNQKNPH---C-----VSLYKLSPPDDPTCK 596
DB 440 TKVTCL-----SCELNPERCOYYSVSFEAKYQLRCSGPGPLPLYTLHSSVNDKGLR 492

QY 597 TKFWMATILDSAGPLPDYTPPE---IFSPESTGTGLYGMLYKPHDLQPKKYPTVLFIY 653
DB 493 VLED-NSALDKM--LQNVQMPKSKLDFFILNETKF-WYQMLPPH-FDKSKKYPILLDVI 547

QY 654 GGPQVQLVNNRFGVKYKFLN---TLASLGYYVVV--IDNRGSHRGKLFEGAFKYMGOI 709
DB 548 AGPCSQ-----KADTVFLRNWATYLASTENIIVASFDGRSGYQGDKIMHAINRRLGTF 601

QY 710 EIDDOVEGLQYLASRYDFDLDRVGTHGWSYGYLSLMAQMORSDFRVAIAGAPVTLWI 769
DB 602 EVEDQIEAARQF--SKMGFVNDKRIATWGSYGYVTSVMVLSGSGVFKCGIAPVSRWE 660

QY 770 FYDTGYTERYMG--HPDQNEQGYLGSVAMQAEKFPSPNRLILLHGFLDENVHFAHTSI 827
DB 661 YDSVYTERYMGUPTEDNLDHYRNSVTMSRAENF--KQVEYLLIHGTADDNVHFQOQSAQ 718

QY 828 LLSFLVRAGKPYDLQIYPOERHSIRVPESGEHYELHLLHYLQE 870
DB 719 ISKALVDVGVDFQAMWYTDHGHCIASSTAHOHYTHMSHFQ 761

RESULT 8

US-11-208-288-2
; Sequence 2, Application US/11208288
; Publication No. US20060051366A1
; GENERAL INFORMATION:
; APPLICANT: CHANG, Chiwen
; TITLE OF INVENTION: USE OF SOLUBLE CD26 AS INHIBITOR OF
; ANGIOGENESIS AND INFLAMMATION
; FILE REFERENCE: 39533-0001
; CURRENT APPLICATION NUMBER: US/11/208,288
; CURRENT FILING DATE: 2005-08-18
; PRIOR APPLICATION NUMBER: US 60/605,013
; PRIOR FILING DATE: 2004-08-26
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Homo sapiens

US-11-208-288-2

Query Match 11.3%; Score 529; DB 7; Length 766;
Best Local Similarity 26.5%; Pred. No. 6.7e-37;
Matches 186; Conservative 109; Mismatches 280; Indels 128; Gaps 32;

QY 206 PADPDMW-----AFHSNDIWSNIVTREERRLYVHNLNMEEDARSAGVATFV 256
DB 149 PNTQWTVSPVGHKLAYWNNDIYVKIEPNLPSYRITWTG-----KEDIYNGITDMV 202

QY 257 LQEE-FDRYSGYWCPCAETTPSGGKILRLIYEENDESEVEIHH---VTSPLMETRRADS 312
DB 203 YEEEFVSAYSALWWSNGTF-----LAVAQNDTEVPLIEYSFYSDLSQVPKTVR 253

QY 313 FRYPKGTGTANPKVTFKMSIIMIDAGRIIDVIDKELIQPFIILFEGVEYIARAGWTPGK 372
DB 254 VPYPKAGAVNPTVKFV--VNTDSLSSVTNATSIQITAPASMLI-GDHYLCDVTWA---- 306

QY 373 YAWSILLDRSQRLQIQLVILSPFLFIPVEDDVMERQRLIESVPDSVTPLLIYEETTDIWIN 432
DB 307 -----TQERISLQWL-----RRIQNY--SYMDICDYDESSGRW-N 338

QY 433 IHDIFHVFPQSHHEEIEFIFASECKTGRH-----LYKITSILKESKYKRSSGGLPAP 485
DB 339 C-----LVARQHIEMSTTGWGRFRPSEPHFTLDGNSFYKIIS--NEEGYRHI----- 384

QY 486 SDFKCPKEIAITSGEWELGRHGSNIQVDEVRRLVYFEGT-KDSPLEHHLVYVSVNP 544
DB 385 CYFQIDKKOCTFITKTWEVIG-----IEALTSYLYIISNEYKMGPGGRNLYKIQLIDY 439

QY 545 GEVTRLDRGSHSCCIS--QHCDFFISKYSNQKNPH---C-----VSLYKLSPPDDPTCK 596
DB 440 TKVTCL-----SCELNPERCOYYSVSFEAKYQLRCSGPGPLPLYTLHSSVNDKGLR 492

QY 597 TKFWMATILDSAGPLPDYTPPE---IFSPESTGTGLYGMLYKPHDLQPKKYPTVLFIY 653
DB 493 VLED-NSALDKM--LQNVQMPKSKLDFFILNETKF-WYQMLPPH-FDKSKKYPILLDVI 547

QY 654 GGPQVQLVNNRFGVKYKFLN---TLASLGYYVVV--IDNRGSHRGKLFEGAFKYMGOI 709
DB 548 AGPCSQ-----KADTVFLRNWATYLASTENIIVASFDGRSGYQGDKIMHAINRRLGTF 601

QY 710 EIDDOVEGLQYLASRYDFDLDRVGTHGWSYGYLSLMAQMORSDFRVAIAGAPVTLWI 769
DB 602 EVEDQIEAARQF--SKMGFVNDKRIATWGSYGYVTSVMVLSGSGVFKCGIAPVSRWE 660

QY 770 FYDTGYTERYMG--HPDQNEQGYLGSVAMQAEKFPSPNRLILLHGFLDENVHFAHTSI 827
DB 661 YDSVYTERYMGUPTEDNLDHYRNSVTMSRAENF--KQVEYLLIHGTADDNVHFQOQSAQ 718

QY 828 LLSFLVRAGKPYDLQIYPOERHSIRVPESGEHYELHLLHYLQE 870
DB 719 ISKALVDVGVDFQAMWYTDHGHCIASSTAHOHYTHMSHFQ 761

RESULT 9

US-10-522-789-2
; Sequence 2, Application US/10522789
; Publication No. US20050260732A1
; GENERAL INFORMATION:
; APPLICANT: TANABE SEIVAKU CO., LTD.
; TITLE OF INVENTION: Three-dimensional structure of dipeptidyl peptidase IV
; FILE REFERENCE: 03-039-PCT
; CURRENT APPLICATION NUMBER: US/10/522,789
; CURRENT FILING DATE: 2005-01-28
; PRIOR APPLICATION NUMBER: US 60/398,761
; PRIOR FILING DATE: 2002-07-29
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Homo sapiens

US-10-522-789-2		ORGANISM: Mus musculus	
US-11-208-288-6		Query Match 11.0%; Score 517.5; DB 7; Length 760;	
		Best Local Similarity 25.0%; Pred. No. 6.5e-36;	
		Matches 186; Conservative 108; Mismatches 277; Indels 136; Gaps 33;	
QY	206	PADPDWI-----AFIHSNDIWSNIVTREERRLTYVHNELANMEEDARSAGVATFV	256
Db	149	PNTQWTVSPGVGHKLAIVVNDIYVKIEPNLPSYRITWTG-----KEDIYINGITDQV	202
QY	257	LOEE--FDYSGVWCPRAETTPSGGKILRLIYEENDESEVEIHH---VTSPLMLETTRADS	312
Db	203	YEEEVFSAYSALWSPNGTF-----LAYAQFNDTEVPLIEYSFSDSLQVPTVR	253
QY	313	FRYKTYGTANKPVFKMSEIMIDAGRIIDVIDKELIQPEILFEGVEYIARAGWTPGK	372
Db	254	VYPKAGAVNTPVAFV--VNTDSLSTVNTATSIQITAPASMLI--GDHYLCDVTWA----	306
QY	373	YAMSTLLDRSOTRQIVLISPELIPVEDDVMERORLIESVPDSVTPLIYEETTDIWIN	432
Db	307	-----TOERISLOWL-----RRIQNY--SVMDICDYDESSGRW-N	338
QY	433	IHDIFHVFPQSHBEIEFIPASECKTGFRH-----LYKITSILKESKYKSSGGLPAP	485
Db	339	C-----LVARQHEINMTGWGRPRPSEPHFTLDGNSFYKITS--NEEGYRHI-----	384
QY	486	SDPKCPKEIAITSGEMWLGHRGNSIQVDEVRRLVYFEGT--KDSPLEHLYVVSYNP	544
Db	385	CYFOIDKKDCTFITKGTWEVIG-----IEALTSDLYIYSNQYKEMPGGRNLYKI----	434
QY	545	GEVTRLDTRGYSHSCCIS-----QHCDFFISKYSNQKNPH---C-----VSLYKLS	592
Db	435	----QLSD--YTKVTCLSCELNPERCQVYSFSKEAKYQLRCGPGPLTYLTHSSVND	488
QY	593	PTCKTKEFWATILDSAGLPDYPPE---IFSPESTTGFTLYGMLYKPHDLQPGKKYPTV	649
Db	489	KGLRVLED--NSALDRW--LQNVQMPESKKLDFILNETKF--WYQMILPPH--FDKSKKYP	543
QY	650	LETYGGPQVLVNNRFGVKYFRLN---TLASLGYYVVV--IDNRGSGHRLKFEAGFYK	705
Db	544	LDVYAGPCSQ-----KADTVFLNWTATYLASTEIIIVASFDGRSGYQGDKIMHAINRR	597
QY	706	MQGIEIDDOVEGLQYASRYDFIDLDRVGTHGWSYGYLSLMALMORSIDFRVATAGAPV	765
Db	598	LGTFFEDDQIEARQF--SKMGFVDNKRIAGWSYGYVTSWVLGSGGVFKGIAVAPV	656
QY	766	TLWIFVDTGYTRYNG--HPQNEQGYLGSVAMQAEKFPSEPNRLLHLHGLDENVHEA	823
Db	657	SRWEYSDSVYTRYMGLPTPEDNLHYNSTVMSRAENP--KQVEYLLHGTADDNVHFQ	714
QY	824	HTSILLSFLVRAGKPYDQIYPOERHSIRVPESGEHVELLHLYQE	870
Db	715	QSAQISKALVDVGVDFQAMWYTDDEHGIASSTAHOHIYTHMSHFQK	761
RESULT 10			
US-11-208-288-6			
; Sequence 6, Application US/11208288			
; Publication No. US20060051366A1			
; GENERAL INFORMATION:			
; APPLICANT: CHANG, Chiwen			
; TITLE OF INVENTION: USE OF SOLUBLE CD26 AS INHIBITOR OF			
; ANGIOGENESIS AND INFLAMMATION			
; FILE REFERENCE: 39533-0001			
; CURRENT APPLICATION NUMBER: US/11/208,288			
; CURRENT FILING DATE: 2005-08-18			
; PRIOR FILING DATE: US 60/605,013			
; NUMBER OF SEQ ID NOS: 8			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 6			
; LENGTH: 760			
; TYPE: PRT			

US-11-208-288-6		ORGANISM: Mus musculus	
		Query Match 11.0%; Score 517.5; DB 7; Length 760;	
		Best Local Similarity 25.0%; Pred. No. 6.5e-36;	
		Matches 203; Conservative 107; Mismatches 306; Indels 197; Gaps 37;	
QY	131	MYSREELLRERKRIGTVGI-----ASYDYHQGGTFLP-----	OAGS 168
Db	67	LYKQENNTLLNAEHNSSIFLENSTFESGYHSVSPDLFVLELNYVYKQWRHSTASY	126
QY	169	GIYHYKDGPGQFTQQLRPNLNVETSCPNIRMPKLCPADPDWI-----AFIHSND	219
Db	127	NIYDYNK-----ROLITEEKIPN-----NTQMITWSPGHKLAYVKNKD	165
QY	220	IMISNIVTREERRLTYVHNELANMEEDARSAGVATFVLOBE--FDYSGVWYWCPC	272
Db	166	IYKVEPHLPASHRIT-----STGEENVYINGITDMWYEEVEFGAYSALWSPNNTFLAY	219
QY	273	AETTPSGGKILRLIYEENDESEVEIHHVTSPLMLETTRADSFRYPKGTGTANPKVTFK	332
Db	220	AQFNDTGVPLEIYSP--YSDS-----LQYPKTVWIPYKAGAVNPTVKFFI--V	265
QY	333	MIDAGRIIDVIDKELIQPEILFEGVEYIARAGWTPGKYA-----WSILLDRSQ	383
Db	266	NIDSLSSSSAAPIQIPAPASVA--RGDHYLCDVVVWATEERISLQWLRRIONYSVMAICDY	324
QY	384	TRLOIVLISPELIPVEDDVMERORLIESVPDSVTPLIYEETTDIWINIHDFHFPQS	443
Db	325	DKINLTWNCPS-----EQQHV-----EMSTTGWVG--RFRPAEPHF	358
QY	444	HEEETFPASECKTGFRHLYKITSILKESKYKSSGGLPAPSPDFK--CPIKEEIAITSGE	502
Db	359	TSOGGSFYKILSDKQGYKHCHF-----PKDKDCTF-----ITKGA	395
QY	503	WEVLGHRGNSIQVDEVRRLVYFEGT--KDSPLEHLYVVSYNPEVTRLDTRGYSHSCCI	561
Db	396	WEVI-----SIALTSDLYIYSNQYKEMPGGRNLYKI-----QLTD--HTNVKCL	439
QY	562	S-----QHCDFFISKYSNQKNPH---C-----VSLYKLS	609
Db	440	SCDLNPERCQYAVSFSKEAKYQLCGWGPGLPLYLHRSTDHKLRLVLE-----DNSA	493
QY	610	---PLPDYTPPE---IFSPESTTGFTLYGMLYKPHDLQPGKKYPTVLFYGGPQVQLVNN	663
Db	494	LDRMLQDVQMPESKKLDFVLNETRF--WYQMILPPH--FDKSKKYPLLLDVYAGPCSQ----	547
QY	664	RFGVKYFRLN---TLASLGYYVVV--IDNRGSGHRLKFEAGFYKMKGOIEIDDOVEGLQ	719
Db	548	--KADASFRLNWTATYLASTEIIIVASFDGRSGYQGDKIMHAINRRLGTLEVEDQIEAAR	605
QY	720	YLASRYDFIDLDRVGTHGWSYGYLSLMALMORSIDFRVATAGAPVTLWIFDTGYTRY	779
Db	606	QFV--KMGFVDSKKVAIWGWSYGYVTSWVLGSGGVFKGIAVAPVSRWEYDSVYTRY	664
QY	780	MG--HPDQNEQGYLGSVAMQAEKFPSEPNRLLHLHGLDENVHFAHTSILLSFLVRAGK	837
Db	665	MGLPIPEDNLHYNSTVMSRAEHF--KQVEYLLHGTADDNVHVFQSAQISKALVDAGV	722
QY	838	PYDQIYPOERHSIRVPESGEHVELLHLYQE	870
Db	723	DFQAMWYTDDEHGIASSTAHOHIYTHMSHFLOQ	755
RESULT 11			
US-11-186-284-55			
; Sequence 55, Application US/11186284			
; Publication No. US20050266493A1			
; GENERAL INFORMATION:			
; APPLICANT: Millennium Pharmaceuticals, Inc.			
; APPLICANT: Berger, Allison			
; APPLICANT: Guillemette, Tracy L.			
; APPLICANT: Kamatkar, Shubhangi			

APPLICANT: Schlegel, Robert
APPLICANT: Monahan, John E.
APPLICANT: Thibodeau, Stephen N.
APPLICANT: Burgat, Lawrence J.
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
TITLE OF INVENTION: THERAPY OF COLON CANCER
FILE REFERENCE: MP01-029P2RN
CURRENT APPLICATION NUMBER: US/11/186,284
CURRENT FILING DATE: 2005-07-21
PRIOR APPLICATION NUMBER: US/10/301,822
PRIOR FILING DATE: 2002-11-21
PRIOR APPLICATION NUMBER: US 60/339,971
PRIOR FILING DATE: 2001-12-10
PRIOR APPLICATION NUMBER: US 60/361,978
PRIOR FILING DATE: 2002-03-05
PRIOR APPLICATION NUMBER: US 60/381,988
PRIOR FILING DATE: 2002-05-20
NUMBER OF SEQ ID NOS: 228
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 55
LENGTH: 760
TYPE: PRT
ORGANISM: Homo Sapiens
US-11-186-284-55

Query Match 9.7%; Score 454; DB 7; Length 760;
Best Local Similarity 24.1%; Pred. No. 1.9e-30; Indels 150; Gaps 27;
Matches 169; Conservative 113; Mismatches 269

Qy 212 IAFHSNDIWSINVTREERLLTYVHNELANEEDARSAGVATFVLQEEF--DRYSGYWM 269
Db 162 LAYVQNNIYLKQPGPPQITF-----NGRENKIFNGIPDWVVEEMLPKYA-LWM 214

Qy 270 CPKAETPSGKILRLIYENDESEVEIHTVSPMLET--RRADSFYPTGTANTKVTFK 328
Db 215 SP-----NGKFL--AYAFNDKXIDPIVAYSYGDEQYPTINIPYKAGANFVARI- 264

Qy 329 MSEIMDAEGRIIDVIDKELQPEI-----LPEGVEYIARAGWTPEGKYANSLDRSQ 383
Db 265 -----FIIDTTPYAVGPOEVPVPMIASSDYFSWLTWTDVRCVLOWI----- 309

Qy 384 TRLQIVLISPELFIPEVDVNERQRLIESVPDSVT-----PLIYEBETTDIWINI 433
Db 310 KVVQNVSVLSCDFREDQWDCPKTOEHIEESRTGWAGGFVSRPVFSYDA-----ISY 364

Qy 434 HDIFHPVQSHHEEIEFIFASECKTGPRHLYKITSILKESKYRSGGLPAISDFKCPK 493
Db 365 VKIF-----SDKDGKHIHI-----KDTVE 385

Qy 494 BEIATSGEWEVLGRHGSNIQVDEVRLLVFECKTOSPLEHLLYVVSYPNPGVETRLTDR 553
Db 386 NAIQITSGKWEAI-----NI-----FRVQDSLFYSSNEFEYPCRRNIYRISIG 430

Qy 554 GYSHS-CCISOH-----CDFEISKYSNKNPHCVSLY-----KLSSPEDDPTCKTEFWAT 603
Db 431 SYPPSKCVTKHLRKCRCQYVTSFSDYAKYALVCVPGIPISLTHDGRTOE-----IK 486

Qy 604 ILDSAGPLPD-----YTPPEIFSPFSTGFTLYGMLYKPHDLQPGKYPVTLVFIYGGPV 658
Db 487 ILEENKELNALKNQLPKBEIKKLEVDEITLWYKMLLPQFDKRSKKYPLLIIQVGGPCS 546

Qy 659 QLVNRFKGVKYFRNLTLASL-----GYVVVIDNRGSGHCHLPEGAFKVGKQTEIDQ 714
Db 547 QSVRS-----VFAVNWISYLSKSGEWMVIALVDGRGTAFQDGLKLLVAVRKUGVYEVDQ 600

Qy 715 VEGLOYLASRYDFIIDLVRGTHGWSYGGYLSLMAIMORSIDIFRVAIAGAPVTLWIFYDTG 774
Db 601 ITAVRKF-EMGFIDEXRIALWHSYGGYSSILASGTGLFKCGIATAVPSSWEYASV 659

Qy 775 YTERYMGHP--DQNEQGYLGSVAMQAEKFPSEPNRLLLHGLFDENVHFAHTSILLSP 832
Db 660 YTERFMGLPTKDDNLEHYKNSTVMAEYFRNVD--YLLIHGTADDNVHFNQNSAQIAKAL 717

Qy 833 VRAGKPYDLQIYQERHSIRVPESG---EHVELLHLHYLOE 870
Db 718 VNAQVDFQAMWYSDQHGL-----SGLSTNHLTYHTHFLKQ 754

RESULT 12
US-11-176-951-11
Sequence 11, Application US/11/176951
Publication No. US20060024313A1
GENERAL INFORMATION:
APPLICANT: CHEN, XIN
APPLICANT: CHEN, YUAN-SHOU
TITLE OF INVENTION: AGENTS THAT DISRUPT DIMER FORMATION IN DPP-IV FAMILY OF
TITLE OF INVENTION: PROLYL DIPEPTIDASES
FILE REFERENCE: 08842.0019
CURRENT APPLICATION NUMBER: US/11/176,951
CURRENT FILING DATE: 2005-07-06
PRIOR APPLICATION NUMBER: 60/586,095
PRIOR FILING DATE: 2004-07-06
PRIOR APPLICATION NUMBER: 60/585,952
PRIOR FILING DATE: 2004-07-06
NUMBER OF SEQ ID NOS: 22
SOFTWARE: PatentIn Ver. 3.3
SEQ ID NO 11
LENGTH: 99
TYPE: PRT
ORGANISM: Homo sapiens
US-11-176-951-11

Query Match 8.4%; Score 394; DB 7; Length 99;
Best Local Similarity 72.7%; Pred. No. 1.2e-26;
Matches 72; Conservative 10; Mismatches 17; Indels 0; Gaps 0;

Qy 774 GYTERYMGHPDQNEQGYLGSVAMQAEKFPSEPNRLLLHGLFDENVHFAHTSILLSP 833
Db 1 GYTERYMGHPDQNEQGYLGSVAMQAEKFPSEPNRLLLHGLFDENVHFAHTSILLSP 60

Qy 834 RAGKPYDLQIYQERHSIRVPESGHEHLLHLYLOENL 872
Db 61 RAGKPYDLQIYQERHSIRVPESGHEHLLHLYLOENL 99

RESULT 13
US-11-179-977-1
Sequence 1, Application US/11/179977
Publication No. US20050249789A1
GENERAL INFORMATION:
APPLICANT: Genencor International, Inc.
TITLE OF INVENTION: Alpha/Beta Hydrolase-Fold Enzymes
FILE REFERENCE: GC511-PCT
CURRENT APPLICATION NUMBER: US/11/179,977
CURRENT FILING DATE: 2005-07-12
NUMBER OF SEQ ID NOS: 21
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1
LENGTH: 657
TYPE: PRT
ORGANISM: Bacillus
US-11-179-977-1

Query Match 5.3%; Score 249.5; DB 7; Length 657;
Best Local Similarity 21.4%; Pred. No. 6e-13;
Matches 150; Conservative 101; Mismatches 265; Indels 185; Gaps 31;

Qy 245 EDARSAGVATFVLQEEFDRYSGYVWCPKAEI-----TPSGGKILRLIL 286
Db 24 DGTAAAYKSVQVQNEK-DSYTSINIYETKGTGGSVPVTHGKRGSTDPKWSPDGRTLAFIS 82

Qy 287 YEENDESEVEIHTVSPMLETRADSFYPTGTANTKVTFKMSEIMID---ABGRIIDV 343
Db 83 DRGDAQAQLYIMSTEGG--EARKLTDIPY---GVSKPLWSPDGESILVTSIGEGESIDD. 137

```
QY 344 IDK---ELIOPFEILFEGVEYIARAGWTPEGKYAMSLIDRSQTRQLQIVL-ISPSELF-----IPVEDD----- 400
Db 138 REKTEQDSYEBEV- -OGLSYKRDGKGLTRGAY-----QLVLVSVK----- 177
QY 401 DVNVERQRLIESVPDSVTPL- -IIVEET---TDIWINIHIDIFHVFPQSHHEEIEF 450
Db 178 -SGEMKELTSHKADHDGDPAPSPDGKWLVPFSAFLTETDDASKPHDVY-----IMS 225
QY 451 IPASECK--TGFRILYKITILKESKYKRSSGGLPAPSPDKPCPIKEEIAITSGEW----- 503
Db 226 LESGDLKQVTPHRGSGFSSFS PDGRYLALLGN- -EKEYK-----NATLSKAWLYDIE 276
QY 504 -----EVLGRH-----GSNIQVDEVRLVYFEGTKDSPLEHHLVYVSVNPGVTR 549
Db 277 QGRLTCLTEMLDVHLADALIGDSLIGGAEQRPW---TKDS---QGFYVIG----- 321
QY 550 LTRG-----YSHSCCISOHCDFEISKYSNOKNPHCVSLYKLSS 588
Db 322 -TDQSTGIYVISTEGLVYFIRLEKEYINSFSLSPDEQHFIAFVTKPDRPSELVSLPQ 380
QY 589 PEDDPTCKTKEFWATILDSAGPLPDYT--PPEIFSEFESTGTGLYMLYKPHDLQPGKY 646
Db 381 BEKQLTGANDKF-----VREHTISIPETIATQEDGVVNGWLMRPAOMEGETTY 430
QY 647 PTVLFYGGPOVLVNNRFXGVYF-RLNLTASLGVVYVVDNRSCHRGKLFEGAFKYY 705
Db 431 PLILNIHGPHMY-----GHTYFHEFQVLAAGVAVYINPRGSHGYGOEFVNAVGRD 484
QY 706 MGQIBIDQVEGLQYLASRYDFIDLDRVGIHGSYGGYLSLMAWMORSDFIRVAIAGAPV 765
Db 485 YGGKDYDDVMOAVDEAIKRDHIDPKRLGVGGSYGGMFTNW-IVGQTNRFKAAVTORSI 543
QY 766 TLWIFY-----DTGY--TERVYNGHP-----DQNEQGYLGSVAMQAEKFPSEPNRL 810
Db 544 SNWISFHGVSIDIGYFFTDWOLEHDMFEDTEKLWDRSPLKY-----AANVE-----TPLL 592
QY 811 LLHGLFDENVHFAHTSILLFLVRAGKPYDLQIYPOQRHSI 851
Db 593 ILHGRDRCPIEQAEQFLTALKKMGKETKLVRFNASHNL 633

RESULT 14
US-11-079-463-7504
; Sequence 7504, Application US/11079463
; Publication No. US20060073161A1
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTEROIDES FR
; FILE REFERENCE: PATH00-03DIV2
; CURRENT APPLICATION NUMBER: US/11/079,463
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/128,705
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: US 09/540,209
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 7504
; LENGTH: 624
; TYPE: PRT
; ORGANISM: B. fragilis
US-11-079-463-7504

Query Match 5.1%; Score 241.5; DB 7; Length 624;
Best Local Similarity 20.6%; Pred. No. 2.7e-12;
Matches 136; Conservative 88; Mismatches 213; Indels 223; Gaps 30;

QY 266 GYWWCPKAETTPSGKILRILY--EENDESEVEIHWTS--MLETRRADSF--RYPKTG 319
Db 77 GFSWSPND-----PLIYYPREGVKDEGPLKRVSPADRPINTRGSRFLARYDIAS 128
QY 320 TANPKVTF-----KXSEIMDAEGRIIDVIDKELIQ-PPEI-----LPEG 358
```

```
Db 129 GTSERLTVGNHSTYMQDISPDGKYLKYSSSKENITQRPFSLSLQVNLTLAVDTLPFE 188
QY 359 VEYIARAGWTPEGKYAMSLIDRSQTRQLQIVL-ISPSELF-----IPVEDD----- 402
Db 189 DRFLGASYSYDPGK-----QLLTASPEAFDGIKGNCONHPIANDFDSQAF 234
QY 403 VNE-QORLIESVPDSVTPLI-----IYBETTDIWINIHIDIFHVFPQSHHEEIEF 450
Db 235 IMDLTRKIDPITKEFNEPSVNFLOWNKDGCYIFSTND-----LNLETDVTS- 272
QY 451 IPASECKTGFRILYKITILKESKYKRSSGGLPAPSPDKPCPIKEEIAITSGEWELGRHG 510
Db 273 ---EDC---RNIRYRSP--KDRKFEK-----LNLETDVTS- 300
QY 511 SNIQVDEVRLVYFEGTKDSPLEHHLVYVSVNPGVTRLTDRGYSHSCCISOHCDFFIS 570
Db 301 -----FAMSENNEP-----SLAAVIGGC----- 318
QY 571 KYSNOKNPHCVSLYKLSSPEDDPTCKTKEFWATILDSAGP-----LPDYTPPEIFSFES 624
Db 319 -----YNAGVAVYDLKK-----KTSRL--IADPMKPTLEKIELGEMKP---WNFTA 360
QY 625 TTGFTLYGMLYKPHDLQPGKYPVLFIYGG--POVOLVNNRFGVKYFELNLTASLYYV 682
Db 361 SDGTEIKGMCLPPSPDPNKYPLIYVYGGTTPTEGINSNYCA-----QLFASRDYV 414
QY 683 VVVDNRSCHRGKLFEGAFKYGQIIBIDQVEGLQYLASRYDFIDLDRVGIHGSYGG 742
Db 415 VVVIQPSGTIGFQGFSAHYNANGKRTADDIIEGTFQCKEHPFVNDKKIGCLGASYGG 474
QY 743 YLSLMAWMORSDFIRVAIAGAPVTLWIFDYGTGYTRYNG-----HPDQNEQGY 791
Db 475 FMT-QYLTQTDIFAAAVSHAGISNVTSY---WGEYGYGYGYNAIAAASYPWNPELFT 530
QY 792 LGSVAMQAEKFPSEPNRLLLHGLFDENVHFAHTSILLFLVRAGKPYDLQIYPOQRHSI 851
Db 531 KQGSLEFNADKINTP---LLHLHGTVDTNVPIGESIQLFNALKILGKTVEFTVDGENHFI 587

RESULT 15
US-11-079-463-9281
; Sequence 9281, Application US/11079463
; Publication No. US20060073161A1
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTEROIDES FR
; FILE REFERENCE: PATH00-03DIV2
; CURRENT APPLICATION NUMBER: US/11/079,463
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/128,705
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: US 09/540,209
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 9281
; LENGTH: 737
; TYPE: PRT
; ORGANISM: B. fragilis
US-11-079-463-9281

Query Match 4.3%; Score 200.5; DB 7; Length 737;
Best Local Similarity 18.8%; Pred. No. 1.2e-08;
Matches 143; Conservative 107; Mismatches 276; Indels 233; Gaps 31;

QY 217 SNDIWTISNIVTREERLT---YVHNE-----LANNEEDARSAGV 252
Db 117 NREVFVNNADGYDNRQIHTHTPYQENEVTVAAADGSKLLFLSNDNGSSQLYEMNPDGSG-- 173
QY 253 ATFVLQEEPDY--SGYWWCPKAETTPSGKILRILYENDESEVEIHWTSPLMLET 308
Db 174 -----RKQISKDYGDIEGY-----SISPDGKILFI-----AQVKTV 205
```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 15, 2006, 00:12:06 ; Search time 11157 Seconds
(without alignments)
4493.674 Million cell updates/sec

Title: US-10-825-632-1

Perfect score: 4700

Sequence: 1 MAAMETFEQLGVEIFETADC.....HLLHYIQENLGRIRALXVI 882

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/abses/ABSSWEB_spool/US10825632/runat_14042006_092559_4657/app_query.fasta.1
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abses08
-USER=US10825632 @CGN 1 1 4939 @runat_14042006_092559_4657 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:*

1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pr:*
9: gb_ro:*
10: gb_sts:*
11: gb_sy:*
12: gb_un:*
13: gb_vi:*
14: gb_htg:*
15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	4700	100.0	2649	6 AX354795	AX354795 Sequence
2	4700	100.0	2656	8 AY172659	AY172659 Homo sapi
3	4700	100.0	2671	6 AR631278	AR631278 Sequence

4	4700	100.0	2671	6 AX608725	AX608725 Sequence
5	4700	100.0	3030	8 AY354202	AY354202 Homo sapi
6	4700	100.0	3106	6 AX342633	AX342633 Sequence
7	4700	100.0	3120	6 AR651452	AR651452 Sequence
8	4700	100.0	3127	6 AF221634	AF221634 Homo sapi
9	4700	100.0	3143	6 AX354793	AX354793 Sequence
10	4696	99.9	4535	8 BC040203	BC040203 Homo sapi
11	4680	99.6	4829	6 AR631283	AR631283 Sequence
12	4680	99.6	4829	6 AX608735	AX608735 Sequence
13	4536.5	96.5	4799	9 BC043124	BC043124 Mus muscu
14	4536.5	96.5	4799	9 BC059222	BC059222 Mus muscu
15	4443.5	94.5	2797	6 AR448400	AR448400 Sequence
16	4397.5	93.6	3125	8 BC030688	BC030688 Homo sapi
17	4385.5	93.3	4685	6 AR631288	AR631288 Sequence
18	4385.5	93.3	4685	6 AX608745	AX608745 Sequence
19	4385.5	93.3	4676	6 AR631287	AR631287 Sequence
20	4385	93.3	4676	6 AX608743	AX608743 Sequence
21	4118	87.6	2842	6 AX405770	AX405770 Sequence
22	4092.5	87.1	4523	6 AR631281	AR631281 Sequence
23	4092.5	87.1	4523	6 AX608731	AX608731 Sequence
24	4091	87.0	2778	8 AX000290	AX000290 Homo sapi
25	3970.5	84.5	2510	6 AX338497	AX338497 Sequence
26	3771	80.2	2668	6 AX405771	AX405771 Sequence
27	3661.5	77.9	4309	6 AR631284	AR631284 Sequence
28	3661.5	77.9	4309	6 AX608737	AX608737 Sequence
29	3364.5	71.6	2161	6 BD157001	BD157001 Primer fo
30	3364.5	71.6	2161	6 AX878058	AX878058 Sequence
31	3364.5	71.6	2161	8 AK027826	AK027826 Homo sapi
32	3103	66.0	2098	6 CQ721272	CQ721272 Sequence
33	2870	61.1	2602	8 AY172660	AY172660 Homo sapi
34	2870	61.1	2617	6 AR631279	AR631279 Sequence
35	2870	61.1	2617	6 AX608727	AX608727 Sequence
36	2870	61.1	3715	8 AK122654	AK122654 Homo sapi
37	2870	61.1	3716	6 AX480934	AX480934 Sequence
38	2870	61.1	4219	6 AR631291	AR631291 Sequence
39	2870	61.1	4219	6 AX608751	AX608751 Sequence
40	2870	61.1	4223	8 BC037948	BC037948 Homo sapi
41	2870	61.1	4295	8 AF452102	AF452102 Homo sapi
42	2870	61.1	4302	6 AR631289	AR631289 Sequence
43	2870	61.1	4302	6 AX608747	AX608747 Sequence
44	2867	61.0	2676	6 AX524928	AX524928 Sequence
45	2863	60.9	2592	8 AY374518	AY374518 Homo sapi

ALIGNMENTS

RESULT 1	AX354795	Sequence 3 from Patent WO0179473.	2649 bp	DNA	linear	PAT 06-FEB-2002
LOCUS	AX354795	Sequence 3 from Patent WO0179473.				
DEFINITION	AX354795					
ACCESSION	AX354795					
VERSION	AX354795.1	GI:18619528				
KEYWORDS						
SOURCE	Homo sapiens (human)					
ORGANISM	Homo sapiens					
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;					
	Homidae; Homo.					

REFERENCE	1	Meyers, R.A. and Williamson, M.
AUTHORS		21953, a human prollyl oligopeptidase family member and uses thereof
TITLE		Patent: WO 0179473-A 3 25-OCT-2001;
JOURNAL		Millennium Pharmaceuticals, Inc. (US)

FEATURES	source	Location/Qualifiers
	1..2649	/organism="Homo sapiens"
		/mol_type="unassigned DNA"
		/db_xref="taxon:9606"

ORIGIN

Alignment Scores:	0	Length:	2649
Pred. No.:	4700.00	Matches:	882

Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	6	Gaps:	0
US-10-825-632-1 (1-882) x AX354795 (1-2649)			
Qy	1	MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys	20
Db	1	ATGCAGCACAATGGAAACAGACAGCTGGGTGGTGGATATTTGAAACTGGCGACTGT	60
Qy	21	GluGlnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr	40
Db	61	GAGGAGAAATTTGAATCAGAGATCGGCTTAATTTGGAGCCCTTTTATGTTGACGGGTAT	120
Qy	41	SerTrpSerGlnLeuLysLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet	60
Db	121	TCCTGGAGTCAGCTTAATAAGCTGCTGCCGATACCGAAATATATCATGGCTACATGATG	180
Qy	61	AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer	80
Db	181	OCTAAGGCACACATGATTTTCATGTTGTGAAGAGGAATGATCCAGATGGACCTCATITCA	240
Qy	81	AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer	100
Db	241	GACAGAATCTATTACCTTGGCATGCTCTGGTGCAACAGAGAAAAATACACTGTTTATCT	300
Qy	101	GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu	120
Db	301	GAATTTCCCAAACTATCAATAGAGCAGCAGTCTTAATGCTCTCTTGGAAAGCCCTCTTTTG	360
Qy	121	AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg	140
Db	361	GATCTTTTTCAGCAACACTGGACTATGGAATGTATTTCTCGAAGAAAGAACTATTAGA	420
Qy	141	GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly	160
Db	421	GNAGAAACCGCATGGACAGTCGGAATTCCTTACGATTATCACCAGGAAGTGA	480
Qy	161	ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly	180
Db	481	ACATTTCTGTTTCAAGCCGTTAGTGGAAATTTATCAGTAAAGATGGAGGGCCACAGGA	540
Qy	181	PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet	200
Db	541	TTTACGCACACACCTTTAAGGCCCATCTAGTGGAACTAGTTGTCCTCCAAACATACGGATG	600
Qy	201	AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle	220
Db	601	GATCCAAAATATTATGCCCTGCTGATCCAGACTGGATTCCTTTATACATAGCAACCATATT	660
Qy	221	TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisGlnLeu	240
Db	661	TGGATATCTTAACATCGTAACCCAGAGAAAGAGAGACTCACTTATGTGCACAAATGAGCTA	720
Qy	241	AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu	260
Db	721	GCCAAACATGGAAAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAAGAA	780
Qy	261	PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly	280
Db	781	TTTGATAGATATCTGGCTATTTGGTGTGTCCTCAAAAGCTCAAACTCCCAAGTGGTGTGT	840
Qy	281	LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal	300
Db	841	AAAAATCTTAGAATTTCTATATGAGAAATGATCAATCTGAGGTGGAAATTTATTCATGTT	900
Qy	301	ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr	320
Db	901	ACATCCCTCTATGTTGGAAACAGAGGGGACAGATTCATTCCTGTTATCTCAAAACAGGTACA	960
Qy	321	AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle	340
Db			
Db	961	GCAAACTCTAAAGTCACCTTTTAAGATGTCAGAAATAATGATGATGCTGAGGAAGGATC	1020
Qy	341	IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu	360
Db	1021	ATGATGTCATAGATAAGGAACATAATTTCAACCTTTTGAATTTCTATTTTGAAGAGTTGAA	1080
Qy	361	TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp	380
Db	1081	TATATTTCCAGAGCTGGATGGACTCTCTGAGGAAAAATATGCTTGGTCCATCTCTACTAGT	1140
Qy	381	ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu	400
Db	1141	CGCTCCACAGACTCGCTTCAGATAGTGTGATCTCACCTGAAATATTTATCTCCAGTAGAA	1200
Qy	401	AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu	420
Db	1201	GATGATGTTATGAAAGGACAGACTCATTTGATGCTGAGTCCCTGATTTCTGTGACGCCACTA	1260
Qy	421	IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe	440
Db	1261	ATTATCTATGAAGAAACCAACAGACATCTGGATAAATATCATGACATCTTTTCATGTTTTT	1320
Qy	441	ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe	460
Db	1321	CCCCAAGTCAAGAGAGGAAATTTGATTTATTTTTCCTCTGAATGCAAAACAGGTTTC	1380
Qy	461	ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly	480
Db	1381	COTCATTTTATCAAAATTTACATCTATTTTAAAGGAAGCAATATAAACGATCCAGTGT	1440
Qy	481	GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer	500
Db	1441	GGGCTGCTGCTCCAGTGATTTCAAGTGCTCTTCAAGAGGAGATAGCAATATCCAGT	1500
Qy	501	GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAsnGluValArgArg	520
Db	1501	GGTGAATGGGAAGTCTTGGCCGCGATGGATCTATATCAAGTTGATGCAAGTCAGAAG	1560
Qy	521	LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisIleLeuTyrValValSer	540
Db	1561	CTGGTATATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACTGTACGTAGTCACT	1620
Qy	541	TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys	560
Db	1621	TACGTAAATCTCTGGAGAGGTGACAAAGCTGACTGACCGTGGCTACTCACATTTCTGCTGC	1680
Qy	561	IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys	580
Db	1681	ATCAGTCAGCACTGTGACTTCTTTTATAAGTAAGTATAGTAACCCAGAGAAATCCACTGT	1740
Qy	581	ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe	600
Db	1741	GTGTCCCTTTCAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAAAACAAAGGAATTT	1800
Qy	601	TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePhe	620
Db	1801	TGGGCCACCAATTTTGGATTACGAGGTCCTCTCTGACTATATCTCTCCAGAAATTTTC	1860
Qy	621	SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu	640
Db	1861	TCCTTTGAAGTACTCTGGATTTTACATTTGATGGATGCTCTACAGGCTCATGATCTA	1920
Qy	641	GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu	660
Db	1921	CAGCTCGAAAGAAATATCTCTACTGTCTGTTCATATATGTTGGTCTCTCAGGTGCGATTG	1980
Qy	661	ValAsnAsnArgPheLysGlyValLysTyrPheArgIleuAsnThrLeuLysSerLeuGly	680
Db	1981	GTGAATTAATCGGTTTAAAGGAGTCAAGTATTTCCGTTTGAATACCTTAGGCTCTTAGGT	2040
Qy	681	TyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly	700
Db	2041	TATGTGTTGTATGATAGACAAACAGGGGATCTCTGTCAACGAGGGCTTAATTTTGAAGGC	2100

```
Qy 701 AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr 720
Db 2101 GCTTTTAAATATAAATAATGGTCAATAGAAATGACCATCAGGTGGAGGACTCCCAATAT 2160

Qy 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrPsrTyr 740
Db 2161 CTAGCTTCTCGATATGATTTTCATTGACTTAGATCGTGGGCATCCACGGCTGCTCTAT 2220

Qy 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
Db 2221 GGAGGATACCTCTCCCTGATGGCATTAAATGCAGAGGTTCAGATATCTTCAGGGTTGCTATT 2280

Qy 761 AlaGlyAlaProValThrIleuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780
Db 2281 GCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACTTATATG 2340

Qy 781 GlyHisProAspGlnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
Db 2341 GGTCAACCTTGACCAATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAA 2400

Qy 801 LysPheProSerGluProAsnArgLeuLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
Db 2401 AGTTCCCTCTGAACCAATCGTTTACTGCTCTTACATGTTTCTTGGATGGAATGTC 2460

Qy 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
Db 2461 CATTTTGCACATACCGATATATCTAGTGTGTTTGTAGTGGGCTGGAAAGCCATATGAT 2520

Qy 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
Db 2521 TTACAGATCTATCTCAGGAGAGACACAGCATAGAGTTCCTGAATCGGAGAACATTAT 2580

Qy 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeuLys 880
Db 2581 GAACTGCATCTTTTGCACCTACCTTCAAGAAACCTTGCATCAGTATTGCTCTAAAAA 2640

Qy 881 ValIle 882
Db 2641 GTGATA 2646
```

```
RESULT 2
AY172659
LOCUS
DEFINITION
Homo sapiens dipeptidyl peptidase IV-related protein-1 (DPRP1)
mRNA, complete cds.
AY172659
AY172659, 1 GI:27549549
```

```
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
```

```
REFERENCE
1 (bases 1 to 2656)
Qi,S., Akinsanya,K., Riviere,P. and Junien,J.-L.
Novel Serine Protease Genes Related To DPPIV
Patent: US (WO 0231134)-A 18-APR-2002;
2 (bases 1 to 2656)
Qi,S., Akinsanya,K., Riviere,P. and Junien,J.-L.
Direct Submission
```

```
JOURNAL
TITILE
Submitted (04-NOV-2002) Perring Research Institute, 3550 General
Atomics Ct., San Diego, CA 92121, USA
Location/Qualifiers
```

```
FEATURES
source
1..2656
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosomes="15"
/map="15q22.1-q22.2"
1..2656
/gene="DPRP1"
8..2656
```

gene

CDS

```
/gene="DPRP1"
/codon_start=1
/product="dipeptidyl peptidase IV-related protein-1"
/protein_id="AA017261.1"
/db_xref="GI:27549550"
/translation="MAAAMETEQLGVETPETAACEENIESQDRPKLEPPFVERYSMSQ
LKKGLADTKYHYGMMAKAPHDPMFKNDPDGPHSDRIYYLAWSGENRENTLIFYSEI
PKTINRAAVLMSWKPLDLFOATLDYGMTSREELLRERKRTGTGVIASDYHQSSG
TLFQAGSGIVHVKGQPGQTPQPLNLTVCNPMRDKPLCPADPDWAFIHSN
DWLSNITRERRLTYVHNELANWEEDASAGVATFVLOEEFDRYGYWCPKAETT
PGGKILRIEENDESEVEIHHVSPMLSTRADSFRYKGTGTANPKVTFKNSLMI
DEAGRIIDVIDKELIQPEILFEGVEYIARAGMTPEKTIWIKWISILLORSQRLQVLIS
PELFIPIEDDVNERQRIETSPVSTPLIIVETTDIWINIHDI FHVFPQSHBEETEF
IFASCKTGFRHLKYITSILKESKVSQSGSLPAPSKPCPKIEEAIATSGEVEILGR
HGSNTQDEVRLRYVPEGKDSPLEHLYVSVYVNGEVTRLTDRGVSQSCCISOHCD
FTISKYSNKNPHCVSLYKLSPEDDPTCKTFWATILDSAGPLDPTTPPEIFSPES
TGTFTLYGMLKPHDLQPKKYPTVFIYGGPQVQVNNRFKGVKFRNLNTLASLYV
VVVDNRGSCHRGLKFEFAGFKYKMGQIEIDQVEGLQYLASRYDFDLDRVGHWSMO
GCYLSLMALMORSDFRVAIAGAPVTLMIFVDTGYTERYMGHPDQNEQGYLGSVAMO
AKKFPSEPNRLLLHLGFLDENHFAHTSILLSLFVRAGKPYDQIYIYQERHSIRVPES
GBHYELHLLHLHQENLGSRIAALKVI"
```

ORIGIN

```
Alignment Scores:
Pred. No.: 0 Length: 2656
Score: 4700.00 Matches: 882
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-825-632-1 (1-882) x AY172659 (1-2656)

Qy 1 MetAlaAlaMetGluThrGluLeuGlyValGluIlePheGluThrAlaAspCys 20
Db 8 ATGCGCAGCAGCAATGGAACAGACAGCTGGGTGTGAGATATTGAAACCTGCGGACTGT 67

Qy 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
Db 68 GAGGAGAATATTGAATCACAGGATCGGCCTAATTTGGAGCCCTTTTATGTTGACGGTAT 127

Qy 41 SerTyrSerGlnLeuLysLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
Db 128 TCTTGAGTCAGCTTAAAGAGCTGCTGCCGATACCAGAAATATCATGGCTACATGATG 187

Qy 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
Db 188 GCTAAGGCACCATCATTTTCATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTTCA 247

Qy 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
Db 248 GACAGAATCTATTACCTTCCCATGTCGTGTGAGAACAGAGAAATATACACTGTTTATCT 307

Qy 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTyrProLysProLeuLeu 120
Db 308 GAAATTTCCCAAACTATCAATAGACAGCAGCTTAATGCTCTCTTGGAGCCCTCTTTTG 367

Qy 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140
Db 368 GATCTTTTTCAGGCAACACTGGACATGTAATGTAATCTTCGAGAGAGAACTATTAAAGA 427

Qy 141 GluArgIleArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
Db 428 GAAAGAAAACGCATTGGAACAGTCGGAATGTCCTTACGATTTATCACCAGGAAGTGA 487

Qy 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
Db 488 ACATTTCTGTTTCAGCCGGTAGTGGATTTATCATCGTAAAGATGGAGGCCACAGGA 547

Qy 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
Db 548 TTACGCAACAACTTTAAGGCCCAATCTAGTGGAACTAGTTGTTGCCCAACATACGCGATG 607
```

Qy	201	AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle	220
Db	608	GATCCAAAATATATGCCCTGCTGATCCAGACTGGATGGCTTTTATACATAGCAACGATATT	667
Qy	221	TrpIleSerAsnIleValThrArgGluGluArgLeuThrTrpValHisAsnGluLeu	240
Db	668	TGGATATCTAACATCGTAACAGAGAAGAAAGAGAGACTCACTTATGTGCACAAATGAGCTA	727
Qy	241	AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu	260
Db	728	GCCAAACATGGAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAAGAAGAA	787
Qy	261	PheAspArgTrpSerGlyTrpTrpCysProLysAlaGluThrThrProSerGlyGly	280
Db	788	TTTGATAGATATCTCGCTATTCGTGTGTCCTCCAAAAGCTGAAACAACTCCCAAGTGGTGT	847
Qy	281	LysIleLeuArgIleLeuTrpGluAsnAspGluSerGluValGluIleIleHisVal	300
Db	848	AAAATTCCTAGATCTTATATGAAGAAATGATGATCTGAGGTGGAAATATTCAATGTT	907
Qy	301	ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTrpProLysThrGlyThr	320
Db	908	ACATCCCTATGTTGGAAACAGAGGGCGAGATTCATTCGTTATCTCTAAACAGGTACA	967
Qy	321	AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle	340
Db	968	GCAAATCTCAAAAGTCATTTTAAAGATGTCAGAAATATGATGATGCTGAGAGAGGATC	1027
Qy	341	IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu	360
Db	1028	ATAGATGTCATAGATAAGGAACATAATTCAACTTTTGAGATCTCTATTTGAAGGAGTTGAA	1087
Qy	361	TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTrpAlaTrpSerIleLeuLeuAsp	380
Db	1088	TATATTGCCAGAGCTGGATGACTCTCTGAGGAGAAATATGCTTGGTCCATCTCTACTAGAT	1147
Qy	381	ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu	400
Db	1148	CGCTCCAGAGCTGCGCTGCAGATAGTGTGATCTCACTCAATATTATTATCCAGTAGAA	1207
Qy	401	AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu	420
Db	1208	GATGATGTTATGGAAGGACAGAGACTCAATTGAGTCAGTCGCTGATTCGTGAGCGCACTA	1267
Qy	421	IleIleTrpGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe	440
Db	1268	ATTATCTATGAAGAAACAAACAGACATCTGGATAAATATCCATGACATCTTTCATGTTTT	1327
Qy	441	ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe	460
Db	1328	CCCCAAAGTCACGAAGAGAAATTCAGTTTATTTTTCCTCTGATTCGAAACAGGTTTC	1387
Qy	461	ArgHisLeuTrpLysIleThrSerIleLeuLysGluSerLysTrpLysArgSerSerGly	480
Db	1388	CGTCATTTATACAAAATTACATCTATTTTAAAGGAAAGCAATATTAACGATCCAGTGT	1447
Qy	481	GlyLeuProAlaProSerAspPheLysCysProIleLysGluIleAlaIleThrSer	500
Db	1448	GGGCTCCCTGCTCCCAAGTGATTTCAAGTGTCCTATCAAGAGAGAGATAGCAATACCACT	1507
Qy	501	GlyGluTrpGluValLeuGluArgHisGlySerAsnIleGlnValAspGluValArgArg	520
Db	1508	GGTGAATGGGAAGTCTTTGGCCGCGCATGGATCTAATATCCAAGTTGATGAAGTCAGAAG	1567
Qy	521	LeuValTrpPheGluGlyThrLysAspSerProLeuGluHisHisLeuTrpValValSer	540
Db	1568	CTGGTATATTTTGAAGGCCACCAAGACTCCCTTTTAGAGCATCACTGTACGTAGTCAGT	1627
Qy	541	TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTrpSerHisSerCysCys	560
Db	1628	TACGTAAATCTGGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCACATTTCTTGCTGC	1687
Qy	561	IleSerGlnHisCysAspPhePheIleSerLysTrpSerAsnGlnLysAsnProHisCys	580
Db	1688	ATCAGTCAGCACTGTGACTTCTTTAAGTAAGTATAGTAACCAAGAAATCCCACTGT	1747
Qy	581	ValSerLeuTrpLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe	600
Db	1748	GTGTCCTTTTACAGCTATCAGTCTCGAAGATGACCCCAACTTGCAAAACAAAGGAATTT	1807
Qy	601	TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTrpThrProGluIlePhe	620
Db	1808	TGGGCCCAACATTTTGGATTTCAGCAGGTCTCTCTCTGACTATACTCTCCAGAAATTTTC	1867
Qy	621	SerPheGluSerThrThrGlyPheThrLeuTrpGlyMetLeuTrpLysProHisAspLeu	640
Db	1868	TCTTTTGAAAGTACTACTCGAATTTACATTTATGAGGATGCTCTACAGCTCATGATCTA	1927
Qy	641	GlnProGlyLysLysTrpProThrValLeuPheIleTrpGlyGlyProGlnValGlnLeu	660
Db	1928	CAGCTCGAAAGAAATATCTACTGTGCTTCATATATGGTGGTCTCAGGTGCAGTGTG	1987
Qy	661	ValAsnAsnArgPheLysGlyValLysTrpPheArgLeuAsnThrLeuAlaSerLeuGly	680
Db	1988	GTGAATATATCGATTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTAGCCCTCTCTAGT	2047
Qy	681	TyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly	700
Db	2048	TATGTGGTGTAGTAGTAGACAACAGGGGATCTGTCCAGAGGGCTTAAATTTTGAAGGC	2107
Qy	701	AlaPheLysTrpLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTrp	720
Db	2108	GCCTTTAAATATAAATGGGTCAAAATAGAAATTTGACGATCAGGTGGAAGACTCCCAATAT	2167
Qy	721	LeuAlaAsnArgTrpAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTrp	740
Db	2168	CTAGCTTCTCGATGATGATTTCAATTGACTTAGATCGTGTGGCATCCACGGCTGGTCTAT	2227
Qy	741	GlyGlyTrpLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle	760
Db	2228	GGAGATACCTCTCCCTGATGGCATTAATGCAAGGTCAGATATCTTCAGGGTTGCTATT	2287
Qy	761	AlaGlyAlaProValThrLeuTrpIlePheTrpAspThrGlyTrpThrGluArgTrpMet	780
Db	2288	GCTGGGGCCCCAGTCACCTCTGGGATCTTCTATGATACAGATACACGGNACGTTATATG	2347
Qy	781	GlyHisProAspGlnAsnGluGlnGlyTrpTrpLeuGlySerValAlaMetGlnAlaGlu	800
Db	2348	GGTCACCTCGACCAAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCCAAGCAA	2407
Qy	801	LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal	820
Db	2408	AAAGTTCCCTCTGAAACCAAAATCGTTTACTGCTCTTACTATGTTTCTCTGGATGAGAATGC	2467
Qy	821	HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTrpAsp	840
Db	2468	CATTTTGCACATACCAGTATATTACTGAGTTTTTTTAGTGAGGGCTGAAAAGCCATATGAT	2527
Qy	841	LeuGlnIleTrpProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTrp	860
Db	2528	TTACAGATCTATCTCTCAGAGAGACACAGCATAGAGTTCTCTGAATCGGAGAACATTTAT	2587
Qy	861	GluLeuHisLeuLeuHisTrpLeuGlnGluAsnLeuGlySerArgIleAlaLeuLys	880
Db	2588	GAATCGATCTTTTGCACCTACCTTCAAGAAAACCTTGGATCAGCTATTGCTGCTCTAAAA	2647
Qy	881	ValIle	882
Db	2648	GTGATA	2653
RESULT 3			
AR631278			
LOCUS			
DEFINITION			
AR631278			
ACCESSION			
AR631278.1			
VERSION			
GI:59770920			
Sequence 2 from patent US 6844180.			
2671 bp			
DNA			
linear			
PAT 14-FEB-2005			

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

Unclassified.

REFERENCE 1 (bases 1 to 2671)

AUTHORS OI, S., Akinsanya, K.O., Riviere, P.J.M. and Junien, J.-L.

TITLE Serine protease genes related to DPPIV

JOURNAL Patent: US 6844180-A 2 18-JAN-2005;

Ferring BV,;

NLX;

FEATURES Location/Qualifiers

source 1..2671

/organism="unknown"

/mol_type="genomic DNA"

ORIGIN

Alignment Scores:

Pred. No.: 0 Length: 2671
 Score: 4700.00 Matches: 882
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 6 Gaps: 0

US-10-825-632-1 (1-882) x ARG31278 (1-2671)

QY 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
 DB 8 ATGGCAGCAGCAGCAATGGAACAGAACAGCTGGTGTGAGATATTTGAACTGGGACTGT 67
 QY 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
 DB 68 GAGGAGATATTTGATACAGATCGGCTTAATTTGAGGCTTTTATGTTGACGGTAT 127
 QY 41 SerTrpSerGlnLeuLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
 DB 128 TCCTGGAGTCAGCTAAAAAGCTGCTTGGCGATACAGAAATATCATGGCTACATGATG 187
 QY 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
 DB 188 GCTAAGGCACACATGATTTTCATGTTGTGAAGAGGAATGATCCAGATGCACCTCATTTCA 247
 QY 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
 DB 248 GACAGAATCTATTACCTTGCCTATGCTGTGAGAACAGAGAAATACACTGTTTATTTCT 307
 QY 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
 DB 308 GAAATTTCCAAACATATCAATAGAGCAGCAGCTCTTAATGCTCTCTTGGAGCCTCTTTTG 367
 QY 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg 140
 DB 368 GATCTTTTTCAGCAACACTGGACTATGGAATGATTTCTCGAGAGAGAACTATTTAAGA 427
 QY 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
 DB 428 GAAAGAAACGCAATGGACACAGTCGGAATGCTCTTACGNTATATCACCAAGGAAGTGA 487
 QY 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
 DB 488 ACATTTCTGTTTCAAGCCGGTAGTGAATTTATCACGTAAAGATGGAGGGCCACAGA 547
 QY 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
 DB 548 TTATCCCAACCACTTTAAGGCCCAATCTAGTGGAACTAGTTGTCCTCCACATACCGATG 607
 QY 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
 DB 608 GATCCAAAATTTATGCCCTGCTGATCCAGACTGGATGCTTTTATACATACACGATATT 667
 QY 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu 240
 DB 668 TGGATATCTAACATCGTAACCAAGAGAAAGAGGAGCTCACTTATGTGCAATGAGCTA 727

QY 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu 260
 DB 728 GCCAACATCGAAGAAGATGCCAGATCAGCTGGAGTCGTACCTTTGTTCTCCAGAAGAA 787
 QY 261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280
 DB 788 TTTGATAGATATTTCTGGCTATTTGGTGTGTCCTCAAAAGCTGAAACAACACTCCCGTGGT 847
 QY 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300
 DB 848 AAAATTTCTTAGAATTTCTATATGAAGAAATGATGAATCTGAGTGGGAAATTTATTCATGTT 907
 QY 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320
 DB 908 ACATCCCTATGTTGGAACAAAGAGGCGCAGATTCATTCCTGTTATCTCTAAACAGGTACA 967
 QY 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
 DB 968 GCMAATCTTAAGTCACTTTTAAGATGTGAGAAATATGATTTGATGTGAGGAGGATC 1027
 QY 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
 DB 1028 ATAGATGTCATAGATAAGGAACCTAATTCACCTTTTGAGATTCATTTTGAAGGAGTTGAA 1087
 QY 361 TyrIleAlaArgAlaGlyTyrThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
 DB 1088 TATATTGCCAGAGCTGGATGGACTCTCGAGGAAATATATGCTGCTCCATCTACTAGAT 1147
 QY 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
 DB 1148 CGCTCCACAGACTCGCTCGCAGATAGTGTGATCTCACCTGAATTTATTTATCCCAAGTAGAA 1207
 QY 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
 DB 1208 GATGATGTTATGGAAGGCGAGACTCATTTGAGTCAGTCGCTGATTTCTGTGACCCACTA 1267
 QY 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
 DB 1268 ATTATCTATGAGAAACAAACAGACATCTGGATTAATATCCATGACATCTTTTCATGTTTTT 1327
 QY 441 ProGlnSerHisGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
 DB 1328 CCCAAAGTCAAGAGAGAAATGAGTTTATTTTGGCTCTGATGCAAAACAGGTTTC 1387
 QY 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
 DB 1388 CGTCATTTATACAAATTTACATCTATTTTAAAGGAAGCAATATATAACGATCCAGTGGT 1447
 QY 481 GlyLeuProAlaProSerAspPheLysCysProLysGluGluIleAlaIleThrSer 500
 DB 1448 GGGCTGCTGCTCCAAAGTGATTTCAAGTGTCTATCAAGAGAGAGATAGCAATTTACCAGT 1507
 QY 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
 DB 1508 GGTGAATGGGAAGTTCTTGGCCGCGATGATCTAATATCCAAAGTTGATGAAGTCAGAAGG 1567
 QY 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer 540
 DB 1568 CTGGTATATTTTGAAGGCGCACCAAGACTCCCCCTTTAGAGCATCACTGTGATGATCAGT 1627
 QY 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
 DB 1628 TAGTAAATCTCGAGAGGTGACAAAGCTGACTGACCGTGGCTACTACATCTTCTGTGTCG 1687
 QY 561 IleSerGlnHisCysAspPheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
 DB 1688 ATCAGTCAAGCACTGTGACTCTTTTATAGTAAGTATAGTAAACAGAGAATCCACACTGT 1747
 QY 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
 DB 1748 GTGTCCCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCCACTTGTGCAAAACAAAGAAATTT 1807

```
Qy 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 620
Db 1808 TGGGCCACCAATTTGGATTTCAGAGGCTCTCTCTCGACTATACCTCCCAAGAAATTTTC 1867
Qy 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
Db 1868 TCTTTTGAAGTACTACTGGATTACATTGTATGGGATGCTCTACAGACCTCATGATCTA 1927
Qy 641 GlnProGlyLysIleValProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 660
Db 1928 CAGCCTGGAAAGAAATATCTACTGTGTCTCATATATGTGTGTCTCAGGTGCAGTTG 1987
Qy 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680
Db 1988 GTGAATATCGAATTAAGGAGTCAAGTATTTTCGGCTTGAAATACCTAGCCTCTAGGT 2047
Qy 691 TyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 700
Db 2048 TATGTGTGTGTAGTGATAGACACAGGGGATCTGTCCAGGAGGCTTAATTTGAGGC 2107
Qy 701 AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr 720
Db 2108 GCCTTTAAATATAAATGGGTCAATAGAAATTTGACGATCAGGTGGAGGACTCCATAT 2167
Qy 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyr 740
Db 2168 CTAGCTTCTCGATATGATTTTCTATGTAGTCTGTGGGCATCCAGGCTGGTCCAT 2227
Qy 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
Db 2228 GGAGGATACCTCTCCCTGATGGCATTAATGCAGAGGTTCAGATATCTTCAGGGTTGCTATT 2287
Qy 761 AlaGlyAlaProValThrLeuTyrIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780
Db 2288 GCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAGCTTATATG 2347
Qy 781 GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
Db 2348 GGTCAACCTGTACCAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAA 2407
Qy 801 LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
Db 2408 AAGTCCCTCTGAACCAATCGTTTACTCTCTTACATGGTTTCTTGGATGAGAATGTC 2467
Qy 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
Db 2468 CATTTTGCATATCCAGTATATTACTGAGTTTTTTAGTGAGGGCTGGAAGGCCATATGAT 2527
Qy 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
Db 2528 TTACAGATCTATCTTCAGGAGACACAGCATAGAGTTCTCTGAATCGGAGAACATTTAT 2587
Qy 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnGlySerArgIleAlaAlaLeuLys 880
Db 2588 GAACTGCACTCTTTTGCACTACCTTCAAGAAACCTTTGGATCAGTATTGCTGCTCTAAA 2647
Qy 881 ValIle 882
Db 2648 GTGATA 2653

RESULT 4
AX608725 LOCUS
DEFINITION Sequence 2 from Patent WO0231134.
ACCESSION AX608725
VERSION AX608725.1 GI:28404298
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominae; Homo.
REFERENCE 1
```

```
AUTHORS Qi,S., Akintanya,K.O., Riviere,P.J. and Junien,J.L.
TITLE Novel serine protease genes related to appiv
PATENT: WO 0231134-A 2 18-APR-2002;
FERRING BV (NL)
FEATURES
    source
        1..2671
        /organism="Homo sapiens"
        /mol_type="unassigned DNA"
        /db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 2671
Score: 4700.00 Matches: 882
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0
US-10-825-632-1 (1-882) x AX608725 (1-2671)
Qy 1 MetAlaAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
Db 8 ATGGCAGCAGCAATGGAAACAGACAGCTGGGTGTGGAGATATTGAAAATCGCGACTGT 67
Qy 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
Db 68 GAGGAGATATTGAATCACAGGATCGGCCTAATTTGGAGCCCTTTTATGTTGAGCGGTAT 127
Qy 41 SerTyrSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
Db 128 TCCTGGATGAGCTTAAAGCTGCTTGCAGTACCCAGAAATATCATGGCTACATGATG 187
Qy 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
Db 188 GCTAAGGACCAACATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGACCTCATTC 247
Qy 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
Db 248 GACAGAACTTATTAACCTTGGCCATGCTGGTGAGAACAGAGAAATACACTGTTTATCT 307
Qy 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTyrLysProLeuLeu 120
Db 308 GAAATTCGCAAAACATCATCATAGACGACAGTCTTAATGCTCTCTTGGAGGCCCTTTTG 367
Qy 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg 140
Db 368 GATCTTTTTCAGGCAACACTGCACTATGGAATGTATTCTCGAGAAGAAGAACTATTAAAG 427
Qy 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
Db 428 GAAAGAAAACGCAATTTGGAAACAGTCGGAATGCTTCTTACGATTATCACCAAGGAAGTGA 487
Qy 161 ThrPheLeuPheGlnAlaGlySerGlyTyrHisValLysAspGlyGlyProGlnGly 180
Db 488 ACATTTCTGTTTCAAGCCGCTAGTGGAAATTTATCAGTAAAGATGAGGGGCCCAAGGA 547
Qy 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
Db 548 TTTTACGCAACAACCTTTTAAAGCCCAATCTAGTGGAAACTAGTTGTCCCAACATACGATG 607
Qy 201 AspProLysLeuCysProAlaAspProAspTyrIleAlaPheIleHisSerAsnAspIle 220
Db 608 GATCCAAAATATGCGCTGCTGATCCAGACTGGATGCTTTTATATACATACACAGATATT 667
Qy 221 TrpIleSerAsnIleValThrArgGluGluArgLysLeuThrTyrValHisAsnGluLeu 240
Db 668 TGGATATCTTAACATCGTAACCCAGAGAAAGAGAGACTCACCTATGTGCAATGAGCTA 727
Qy 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu 260
Db 728 GCCAACATGGAAGAAGATGCCAGATCAGCTGAGTGCCTACCTTTGTTTCTCCAGAAAGAA 787
```


Qy	261	PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly	280	Db	1868	TCTTTTGAAGTACTACTGGATTACATTGTATGGGATGCTCTACAGGCTCATGATCTA	1927
Db	788	TTTGATAGATATTCTGGCTATTGGTGTCTCCAAAGCTGAAACAACTCCAGGTGGT	847	Qy	641	GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu	660
Qy	281	LysIleLeuArgIleLeuTyrGluGluAenAspGluSerGluValGluIleHisVal	300	Db	1928	CAGCCTGGAAGAAATATCTCTACTGTCTGTTCATATATGGTGGTCTCCTCAGGTGCAGTTG	1987
Db	848	AAAATTCTTAGAATTTCTATATGAAGAAATATGATGATCTGAGTGGAAATATTATCATGTT	907	Qy	661	ValAenAenArgPheLysGlyValLysTyrPheArgLeuAenThrLeuAlaSerLeuGly	680
Qy	301	ThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThr	320	Db	1988	GTGAATAATTCGATTAAAGGAGTCAGTATTTCCGCTTGAATACCCCTAGGCTCTCTAGGT	2047
Db	908	ACATCCCTATGTTGGMAACAGAGGCGAGATTCAATCCGTTATCTCTAAACAGGTACA	967	Qy	681	TyrValValValValLleAspAenArgGlySerCysHisArgGlyLeuLysPheGluGly	700
Qy	321	AlaAenProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle	340	Db	2048	TATGTGGTGTAGTGATAGACACACAGGGATCTGTCTACCGAGGGCTTAAATTTGAAGGC	2107
Db	968	GCAATCTCTAAAGTCACTTTTAAAGATGTCTAGAAATATGATTGATGCTCAAGGAGATC	1027	Qy	701	AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr	720
Qy	341	IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluValGlu	360	Db	2108	GCCTTTAAATATAAATGGGTCAAAATAGAAATTCAGCATCAGGTGAAGGACTCCAAATAT	2167
Db	1028	ATAGATGTCATAGATAAGGAACATAATTCACACCTTTTGAGATTCTATTTTGAAGGAGTTGAA	1087	Qy	721	LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyr	740
Qy	361	TyrIleAlaArgAlaGlyTyrThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp	380	Db	2168	CTAGCTTCTCGATATGATTTTCATTGACTTAGATCGTGTGGCATCCACGGCTGGTCTAT	2227
Db	1088	TATATTGCCAGACTGGATGGACTCTCTGAGGNAATATGCTGGTCCATCTCTACTAGAT	1147	Qy	741	GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle	760
Qy	381	ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu	400	Db	2228	GGAGGATACCTCTCCCTGATGGCAATTAATGCAGAGGTTCAGATATCTTCAGGGTGTCTATT	2287
Db	1148	CGCTCCAGACTCGCCTGCAGATAGTGTGGTATCTACCTGGAATATTATTATCCAGTAGAA	1207	Qy	761	AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet	780
Qy	401	AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu	420	Db	2288	GCTGGGGCCCCAGTCACCTCTGTGGATCTTCTATGATACAGATACACGGAACGTATATG	2347
Db	1208	GATGATGTTATGGAAGGCGAGACTCATTGAGTCACTGCTCGATTCCTGTGAGCGCACTA	1267	Qy	781	GlyHisProAspGlnAenGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu	800
Qy	421	IleIleTyrGluGluThrThrAspIleTrpIleAenIleHisAspIlePheHisValPhe	440	Db	2348	GGTCAACCTCGACAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGACAGAA	2407
Db	1268	ATTATCTATGAAAGAAACACAGACATCTGGATTAATATCCATGACATCTTTCATGTTTT	1327	Qy	801	LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal	820
Qy	441	ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe	460	Db	2408	AAGTCCCCCTCTGAACCAAAATCGTTTACTGCTCTTACATGGTTCTCTGGATGAGATGTC	2467
Db	1328	CCCCAAAGTCAGGAAGGAAATTTAGTTTATTTTGGCTCTGAAATGCAAAACAGGTTTC	1387	Qy	821	HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp	840
Qy	461	ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly	480	Db	2468	CATTTTGCACATACCAGTATATTACTAGATTTTTTAGTGGGCTGGAAAGCCATATGAT	2527
Db	1388	CGTCATTTATACAAATTTACATCTATTATTTAAAGGAAGCAAAATATAAACGATCCAGTGGT	1447	Qy	841	LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr	860
Qy	481	GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer	500	Db	2528	TTACAGATCTATCTCTCAGGAGACACACATAGAGTTCCTGAATCGGAGAACATTAT	2587
Db	1448	GGGCTGCCTGCTCCAAAGTGATTTCCAGTGCTCTATCAAGAGAGATAGCAATACCACT	1507	Qy	861	GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLys	880
Qy	501	GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg	520	Db	2588	GAACTGCATCTTTTGGCACTTCCAGAAAACCTTCAGAACCTGATCAGTATGTGCTCTAAAA	2647
Db	1508	GGTGNATGGGAAGTTCTTGGCCGCGCATGGATCTAATATCAAGTTGATGAAGTCAGAAG	1567	Qy	881	ValIle 882	
Qy	521	LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisIleHisLeuTyrValValSer	540	Db	2648	GTGATA 2653	
Db	1568	CTGGTATATTTTGAAGGCACCAAGACTCCGCTTTTAGAGCATCACCTGTACGTAGTCACT	1627	RESULT 5			
Qy	541	TyrValAenProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys	560	AY354202			
Db	1628	TACGTAAATCTCGAGAGGTGACAAAGGCTGACCGGTGGTACTACATCTTCTGCTGC	1687	LOCUS			
Qy	561	IleSerGlnHisCysAspPheIleSerLysTyrSerAsnGlnLysAenProHisCys	580	DEFINITION			
Db	1688	ATCAGTCAGCACTGTGACTTCTTTAAGTAAGTATAGTAAACCAAGAAATCCACACTGT	1747	ACCSSION			
Qy	581	ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe	600	VERSION			
Db	1748	GTGTCCCTTACAGCTATCAAGTCTCTGAAGATGACCCCACTTGCAGAAACAAAGGAATTT	1807	KEYWORDS			
Qy	601	TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe	620	SOURCE			
Db	1808	TGGGCGACCATTTTGGATTCTACGAGGTCCTCTTCCTGACTATATCTCTCCAGAAATTTTC	1867	ORGANISM			
Qy	621	SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu	640	Homo sapiens (human)			
				Homo sapiens			
				Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
				Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;			
				Hominidae; Homo.			
				1 (bases 1 to 3030)			
				Sha, J. H., Zhou, Z. M. and Li, J. M.			
				Direct Submission			
				Submitted (29-JUL-2003) Key Lab of Reproductive Medicine, Nanjing			
				Medical University, 140 Han Zhong Road, Nanjing, Jiangsu 210029,			

QY	141	GluArgLysArgGileGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly	160
Db	629	GAAAGAAAAAGCATTTGGAACAGTCGGAATTTGCTTTTACGATTTATCACCAAGAAAGTGA	688
QY	161	ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValValAspGlyGlyProGlnGly	180
Db	689	ACATTTCTGTTTCAAGCGCGTAGTGGAAATTTATCAGCTAAAGATGGAGGGCCACAAAGA	748
QY	181	PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet	200
Db	749	TTTACGCAACAACTTTAAGCCCAATCTAGTGGAAACAGTAGTTGTCCCAACATACGATG	808
QY	201	AspProLysLeuCysProAlaAspProAspTyrPheAlaPheIleHisSerIleAspIle	220
Db	809	GATCCAAATTTATGCCCTGCTGATCCAGATGGATTTGCTTTATATCATAGCAACGATATT	868
QY	221	TyrPheSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu	240
Db	869	TGGATATCTAAACATCGTAACAGAGAAAGAGAGACTCATTATGTGCACAATGAGCTA	928
QY	241	AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu	260
Db	929	GCCAACATGGAAGAGATGCCAGATCAGCTGGAGTCCGCTACCTTTGCTCCAAAGAGAA	988
QY	261	PheAspArgTyrSerGlyTyrTyrTyrCysProLysAlaGluThrThrProSerGlyGly	280
Db	989	TTTGATAGATATTCGGCTATTCGGTGTGCCAAAAGCTGAAAACAACTCCCAAGTGGTGT	1048
QY	281	LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal	300
Db	1049	AAAATCTTAGAATTTCTATATGAAGAAATGATGATGATGAGTGGAAATTTATTCATGTT	1108
QY	301	ThrSerProMetLeuGluThrArgGluAlaAspSerPheArgTyrProLysThrGlyThr	320
Db	1109	ACATCCCTATGTTGGAAAACAAGAGGCGCAGATTCATTCCGTTATCTTAAACAGGTACA	1168
QY	321	AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle	340
Db	1169	GCAAAATCTCAAGTCACCTTTTAAGATGTCAGAAATATGATGATGCTGAGAGAGGATC	1228
QY	341	IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu	360
Db	1229	ATAGATGTCATAGATAAGGAACATAATTCACCTTTTGAGATTCATTTGAAGGAGTTGAA	1288
QY	361	TyrIleAlaArgAlaGlyTyrThrProGluGlyLysTyrAlaTyrPheSerIleLeuLeuAsp	380
Db	1289	TATATTGCCAGAGCTGGATGGACTCTCTGAGGAGAAATATGCTTGGTCCATCTCTACTAGAT	1348
QY	381	ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu	400
Db	1349	CGCTCCAGAGCTCGCCTACAGATAGTGTGATCTCACTCACTGAATTTATTTATCCAGTAGAA	1408
QY	401	AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu	420
Db	1409	GATGATGTTATGGAAGCGCAGAGACTCATTGAGTCACTGCTGATCTGTGACGCCACTA	1468
QY	421	IleIleTyrGluGluThrThrAspIleTyrPheAsnIleHisAspIlePheHisValPhe	440
Db	1469	ATTATCTATGGAAGAAACAACAGACATCTGGATAAATATCCATGACATCTTTCATGTTTTT	1528
QY	441	ProGlnSerHisGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe	460
Db	1529	CCCCAAGTCACGAAGAGAAATTTGAGTTTATTTTTTGCCTCTGATGCAAAACAGGTTTC	1588
QY	461	ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly	480
Db	1589	CGTCATTTATACAAATTTACATCTATTTTAAAGGAAGCAAAATATAAGCATCCAGTGTGT	1648
QY	481	GlyLeuProAlaProSerAspPheLysCysProIleLysGluIleAlaIleThrSer	500
Db	1649	GGGCTGCTGCTCCAGGTGATTTTCAAGTGTCTTATCAAGAGAGAGATGCAATTTACAGT	1708
QY	501	GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg	520

China	Location/Qualifiers
gene	1. .3030 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /chromosome="15" /tissue_type="testis"
CDS	1. .3030 /gene="Dpp8" 161. .2857 /gene="Dpp8" /note="alternatively spliced" /codon_start=1 /product="dipeptidyl peptidase 8 isoform 3" /protein_id="AAQ63887.1" /db_xref="GI:34329352"
ORIGIN	translation="MWKRSEQMKIKSGCKNMAAMETEOLGVEIPETADCNIESOD RPLKPFVVERYSOLKLLADTKYHGYMMAKAPHDFMFKVRNDPDGPRSDRIYLL AMSGENRNTLPYSEIPTKINRAAGLVMLSKPLLDLPQATLDYGMVSRREELLRRK IGTVGIASDYHQSGGTFLOAGSGIVHKVKGQPGQFTQOPLRPNLVRTSPNLRMDP KLCPADPMWIAFTNSNDIMLSNIVTREERLTYYVHNELANMEEDARSAGVATFVLQER PRYSGVMCPKAPFTSGGSKILRLIYEENDESEVERLHVTSPLMETLRADSPFRPKT CTANPKVTFKSEMTMDAGEIITDVIDKELLOPFIPLFGVEYIARAGWPRGKYANS LILDSROELOIVLSELPIETDVEDVMESORLIESVPDSVTPLLIYSEITDINIHL DIPHPQSHSEERIEPASECTGFPHLYKITSILKESKYRSGSGGIPAPSDPKCPI KEBIAITSGEWSICWICDFI5KTSKNQPHCVSYIKDUSPEDDFICTKEFWATILDSA TDGSHSCSSQSGEDFISLWALMQRSDFIRVALAGAPVTLMFIPDYGTERFMGH GKPYFLNTLPSFTGTVVVDINRGSCHRGKFGFKYKMGQIEIDQVQVQLVNNRFK YDFILDRVLIHWSYGYSLSLMALMQRSDFIRVALAGAPVTLMFIPDYGTERFMGH PONSEGVLGVSAMQAEKFPSEPNRLLLHLGFLDENVHPAHTSILLISFLVRAGRPYD LQIYPOERHSIRVPSGHEVELHLLHYLQENLGSRIAALKVI"
Alignment Scores:	
Pred. No.:	0
Score:	4700.00
Percent Similarity:	100.0%
Best Local Similarity:	100.0%
Query Match:	100.0%
DB:	8
US-10-825-632-1 (1-882) x AY354202 (1-3030)	
QY	1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
Db	209 ATGGCAGCAGCAATGAAACAGACAGCTGGGTGTTTCAGATATTTTCAAACCTCGCGAGTGT 268
QY	21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
Db	269 GAGGAGAAATTTGAATCACAGGATCGGCTAAATTTGAGGCTTTTATTTGTTGAGCGGTAT 328
QY	41 SerTrpSerGlnLeuLysIleValLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
Db	329 TCTGAGTGCAGCTTAAAGCTGCTGCCGATACCAAGAAATATCATGGCTACATGATG 388
QY	61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
Db	389 GCTAAGGCACCATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGGACTCATCA 448
QY	81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
Db	449 GACAGAAATCTATTACCTTGCCATGCTGGTGAGAACACAGAAATAACACTGTTTATTCT 508
QY	101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
Db	509 GAAATTTCCCAAACTATCAATAGACGACAGCTCTTAATGCTCTCTTGGAGGCTCTTTTG 568
QY	121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuArg 140
Db	569 GATCTCTTTTCAGGCAACACTGCAGTATCGAATGTATTTCTCGAAGAAAGAACTATTAA 628

```
Db 1709 GGTGAATGGGAAGTTCTTGGCCGCATGGATCTAATATCCAGTTGATCAAGTCAGAAGG 1768
Qy 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSer 540
Db 1769 CTGGTATATTTTGAAGGCACCAAGACTCCCTTTTATAGCATCACCTGTACGTAGTCAGT 1828
Qy 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
Db 1829 TACGTAAATCTCGAGAGGTGACAAAGCTGACTGACGCTGGCTACTACATCTTCTGCTGC 1888
Qy 561 IleSerGlnHisCysAspPheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
Db 1889 ATCAGTCAGCACTGTGACTCTTTATTAAGTAAAGTATAGTAAACCAGGAAGATCCACACTGT 1948
Qy 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
Db 1949 GTGTCCCTTACAGCTATCAAGTCTCTGAAGATGACCCAACTTGCACAAACAAAGGAATTT 2008
Qy 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGlnLysPhe 620
Db 2009 TGGGCACCAATTTTGGATTACAGAGGTCTCTCTCTGACTATACTCTCCAGAAATTTTC 2068
Qy 621 SerPheGluSerThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
Db 2069 TCTTTTGAAGTACTACTGGATTTACATTTGTATGGATGCTCTACAAAGCTCATGATCTA 2128
Qy 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 660
Db 2129 CAGCCTGGAAAGAAATATCTTACTGTCTGTCTCATATATGTTGGTCTCTCAGGTGCAGTTG 2188
Qy 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680
Db 2189 GTGAATAATCGGTTTAAAGAGTCAAGTATTTCCGTTTGAATACCTTAGCTCTCTAGGT 2248
Qy 681 TyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 700
Db 2249 TATGTGGTTGTAGTGATAGACACAGGGGATCTGTCCACGAGGCTTAAATTTGAAGGC 2308
Qy 701 AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr 720
Db 2309 GCCTTTAATATAAATGGGTCAAAATAGAAATTCAGCATCAGGTGAAGGACTCCCAATAT 2368
Qy 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrPheTyr 740
Db 2369 CTAGCTTCTCGATATGATTTCTTCACTTAGATCGTGTGGGCATCCACGCGTGTCTAT 2428
Qy 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
Db 2429 CGAGGATACCTCTCCCTCATGGCATTAATGCAGAGGTTCAGATATCTTCAGGGTTGCTATT 2488
Qy 761 AlaGlyAlaProValThrLeuTyrIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780
Db 2489 GCTGGGGCCCCAGTCACCTCTGTGATCTTCTATGATACAGGATACAGGAACGTTATATG 2548
Qy 781 GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
Db 2549 GGTCAACCTTGACCAAGTAAACAGGGCTATTACTTAGGATCTGTGGCCATGCAGCGA 2608
Qy 801 LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
Db 2609 AAGTTCCTCTCGAACCAATCGTTTACTGCTCTTACATGGTTTCTGGATGAGATGTC 2668
Qy 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
Db 2669 CATTTTGCATACACAGTATATTAAGTTTTTTTGTAGGGCTGGAAGCCATATGAT 2728
Qy 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
Db 2729 TTACAGATCTATCTCTCAGAGAGACACAGCATAGAGTTCTCGAATCGGGAGAACATTAT 2788
Qy 861 GluLeuHisLeuLeuHisTyrLeuGlnAsnLeuGlySerArgIleAlaLeuLys 880
```

```
Db 2789 GAACTCGATCTTTTGCACACTACCTTCAGAAAAACCTTGCATCAGTATTGCTGCTTAAAA 2848
Qy 881 ValIle 882
Db 2849 GTGATA 2854

RESULT 6
AX342633 3106 bp DNA linear PAT 12-JAN-2002
LOCUS Sequence 30 from Patent WO0198468.
DEFINITION AX342633
ACCESSION AX342633
VERSION AX342633.1 GI:18152030
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1
AUTHORS Yue,H., Elliott,V.S., Gandhi,A.R., Lal,P., Au-Young,J.,
Tribouley,C.M., Deleageane,A.M., Baughn,M.R., Nguyen,D.B., Lee,E.A.,
Hafalia,A., Khan,F.A., Walia,N.K., Yao,M.G., Lu,D.A., Patterson,C.,
Tang,Y.T., Walsh,R.T., Azimzai,Y., Ramkumar,J., Xu,Y. and Reddy,R.
JOURNAL Patent: WO 0198468-A 30 27-DEC-2001;
Incyte Genomics, Inc. (US)
FEATURES
source
1..3106
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="incyte ID No: 7160544CBI"

ORIGIN
Alignment Scores: 0 Length: 3106
Pred. No.: 4700.00 Matches: 882
Score: 100.0% Conservative: 0
Percent Similarity: 100.0% Mismatches: 0
Best Local Similarity: 100.0% Indels: 0
Query Match: 6 Gaps: 0
DB:

US-10-825-632-1 (1-882) x AX342633 (1-3106)
Qy 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
Db 203 ATGCGCAGCAGCAATGGAAACAGAACAGCTGGGTGTGAGATATTGAAACCTGGCGACTGT 262
Qy 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
Db 263 GAGGAATATTGAATTCACAGGATCGGCTTAATTTGGAGCCCTTTTATGTTGACGGGTAT 322
Qy 41 SerTyrSerGlnLeuLysLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
Db 323 TCTGTGAGTACGCTTAAAGAGCTGTTCGCGATACCAGAAATATCATGGCTACATGATG 382
Qy 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
Db 383 GCTAAGCCACCATGATTTTCATGTTTGTGAAGGAATGATCCAGATGGACCTCAATCA 442
Qy 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
Db 443 GACAGAATCTATTACCTGCCATGTCGTGTGAGAACAGAGAAATACACTGTTTATTCT 502
Qy 101 GluIleProLysThrIleAsnArgAlaValLeuMetLeuSerTyrLysProLeuLeu 120
Db 503 GAAATTCCTCCAAAATCATCAATAGCAGCAGCTTTAATGCTCTCTTGGAGCCCTCTTTTG 562
Qy 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuArg 140
Db 563 GATCTTTTTCAGGCAACACTGGACTATGGAATGTATTCTCAGAGAAGAACTATTAGA 622
Qy 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
```

Db	623	GAAGAAGAAACGCATTGGACAGCTCGGAATTGCTTCTTACGATTATTCACCAAGGAAGTGGG	521	LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSer	540
Qy	161	ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly	1763	CTGGTATATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTACGTAGTCACT	1822
Db	683	ACATTTCTGTTTCAAGCCGGTAGTGAAATTTATCAGTAAAGATGGAGGGCCCAAGGA	541	TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys	560
Qy	181	PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet	1823	TACGTAAATCTCTGAGAGGTGACAAAGCTGACTGACCGTGGCTACTCACATTTCTTGCTGC	1882
Db	743	TTTACGCAACAACCTTTAAGGCCCAATCTAGTGGAAACTPAGTTGTGCCCAACATACGGATG	561	IleSerGlnHisCysAspPhePheIleSerIleTyrSerAsnGlnLysAsnProHisCys	580
Qy	201	AspProLysLeuCysProAlaAspProAspTrrPileAlaPheIleHisSerAsnAspIle	1883	ATCAGTCAGCACTGTGACTTCTTTATTAAGTAAGTATAGTAACCAAGAAATCCCACTGT	1942
Db	803	GATCCAAAATATATGCCCTGCTGATCCAGACTGGATTGCTTTTATACATAGCAACGATATT	581	ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe	600
Qy	221	TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu	1943	GTGTCCCTTTTCAAGCTATCAAGTCTTCAAGAGTGACCCAACTTGCAAAACAAGGAATTT	2002
Db	863	TGGATATCTAACATCGTAACACAGAGAAGAAAGAGACTCACTATATGTGCACAATGAGCTA	601	TrrAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe	620
Qy	241	AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu	2003	TGGGCCACCAATTTTGGATTTCAGCAGGTCTCTTCTGACTATATCTCTCCAGAAATTTTC	2062
Db	923	GCCAAACATGGAAGAAGATGCCAGATCAGCTGGAGTGCCTTACCTTTGTTCTCCAAAGAAA	621	SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu	640
Qy	261	PheAspArgTyrSerGlyTyrTrrTrrCysProLysAlaGluThrThrProSerGlyGly	2063	TCTTTTGAAGTACTACTTGGATTTACATTTGTATGGGATGCTCTACAAGCCTCATGATCTA	2122
Db	983	TTTGTATAGATATCTGGCTATTTGGTGGTGTCCAAAGCTGAAACAACCTCCCGAGTGGT	641	GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu	660
Qy	281	LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal	2123	CAGCCTGGAAAAAATATCTTCTGTCTATATATGCTGTCTCTCAGGTGCAGTTG	2182
Db	1043	AAAATTTCTAGAAATCTATATGAAGAAATATGTAATCTGAGTGGAAATATATCAATGTT	661	ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly	680
Qy	301	ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr	2183	GTGAATATTCGGTTTAAAGGATCAAGTAATTCGGCTTGAATACCTTAGCTCTCTAGGT	2242
Db	1103	ACATCCCCATGTTTGGAAACAAGAGGGCAGATTCAATCCGTATATCTTAAACAAGGTACA	681	TyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly	700
Qy	321	AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle	2243	TATGTGGTTGTAGTGATAGACACAGGGGATCTCTGTCCACCAGGGCTTAAATTTGAAGGC	2302
Db	1163	GCAAAATCTAAAGTCACTTTTAAAGATGTCAAGAAATATGAATTCATGCTGGAAGAGGATC	701	AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr	720
Qy	341	IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu	2303	GGCTTTAAATATAAATGGGTCAATAGAAATTCAGCTCAGGTGGAGACTCCAATAT	2362
Db	1223	ATAGATGTCAATAGAAAGAACTAAATCAACCTTTTGGAGATCTATTTGAAGAGGTTGAA	721	LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrrSerTyr	740
Qy	361	TrrIleAlaArgAlaGlyTrrThrProGluGlyLysTyrAlaTrrSerIleLeuLeuAsp	2363	CTAGCTTCTCGATATGATTTTCATTGACTTAGATCGTGGGCATCCACGGCTGGTCTAT	2422
Db	1283	TATATTGCCAGAGCTGGATGGACTCTCTGAGGGAAAAATATGCTTGGTCCATCTCTAGAT	741	GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle	760
Qy	381	ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu	2423	GGAGGATACCTCTCCCTGATGGCATTAATGCAGAGGTCAAGATATCTTCAGGGTTGCTATT	2482
Db	1343	CGCTCCAGACTCGCCTACAGATAGTGTTCATCTCACCTGAAATTTATATCCAGTAGAA	761	AlaGlyAlaProValThrLeuTrrPilePheTyrAspThrGlyTyrThrGluArgTyrMet	780
Qy	401	AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu	2483	GCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACGTTATATG	2542
Db	1403	GATGATGTTATGGAAAGGCAGAGACTCAATTSAGTCAGTGCCTGATTCGTGACGCCACTA	781	GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu	800
Qy	421	IleIleTyrGluGluThrThrAspIleTrrPileAsnIleHisAspIlePheHisValPhe	2543	GGTCACCTTGACCAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAA	2602
Db	1463	ATTATCTATGAAGAAACACAGACATCTGGATTAATATCCATGACATCTTTCATGTTTTT	801	LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal	820
Qy	441	ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe	2603	AGTTTCCCCCTCGAACCATAATCGTTTACTGTCTTACATGGTTTCTCGGATGAGAATGTC	2662
Db	1523	CCCCAAAGTCAAGAGGAAATTTAGTTTTATTTTTGGCTCTGAATTCGAAAAACAGGTTTC	821	HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp	840
Qy	461	ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTrrLysArgSerSerGly	2663	CATTTTGCACATACCAGTATATTTACTGATTTTTTTAGTGGGCTGGAAGCCATATGAT	2722
Db	1583	CGTCATTTTATACAAAATTTACATCTATTTTAAAGGAAGCAAAATATATAACGATCCAGTGGT	841	LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr	860
Qy	481	GlyLeuProAlaProSerAspPheLysCysProIleLysGluIleAlaIleThrSer	2723	TTACAGATCTATCTCCAGAGACACAGCATAGAGTTCTGTAATCGGAGAACATATAT	2782
Db	1643	GGGCTGCCCTGCCAAGTGATTTCAAGTGTCTTCAAGAGAGAGATAGCAATATCCAGT	861	GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLys	880
Qy	501	GlyGluTrrPgluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg	2783	GAACGTGATCTTTTGCACACTACCTTCAAGAAAAACCTTGGATCACCATTTGCTGCTCTANAA	2842
Db	1703	GGTGAATGGGAAGTTCTTTGGCCGGCATGGATCTTAATATCCAAGTTGATGAAGTCAGAAAG			

```

Qy      881 ValIle 882
Db      2843 GTGATA 2848

RESULT 7
AR651452
LOCUS      AR651452      3120 bp      DNA      linear      PAT 20-APR-2005
DEFINITION Sequence 2 from patent US 6881564.
ACCESSION AR651452
VERSION    AR651452.1 GI:62795938
KEYWORDS   Unknown.
SOURCE     Unknown.
ORGANISM   Unclassified.

REFERENCE
  1 (bases 1 to 3120)
AUTHORS    Abbott, C.A. and Gorrell, M.D.
TITLE      Dipeptidyl peptidases
JOURNAL    Patent: US 6881564-A 2 19-APR-2005;
           The University of Sydney; Sydney;
           AUX;

FEATURES
  source
  1..3120
  /organism="unknown"
  /mol_type="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.:      0      Length:      3120
Score:          4700.00      Matches:      882
Percent Similarity: 100.0%      Conservative: 0
Best Local Similarity: 100.0%      Mismatches: 0
Query Match:      100.0%      Indels:      0
DB:              6      Gaps:      0

US-10-825-632-1 (1-882) x AR651452 (1-3120)
Qy      1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
Db      214 ATGGCAGCAGCAATGGAACAGACAGCTGGGTGTGAGATATTGGAATCTGGCGACTGT 273
Qy      21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
Db      274 GAGGAGAAATTAATACAGGATCGGCTAAATTTGGAGCCCTTTTATGTGAGCGGTAT 333
Qy      41 SerTrpSerGlnLeuLysLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
Db      334 TCTGGAGTCAGCTTAAAGCTGCTTCCGATACCAAGAAATATCATGCTACATGATG 393
Qy      61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
Db      394 GCTAAGGCACCAATGATTTTCATGTTGTGAAGAGGAATGATCCAGATGGACCTCATTTCA 453
Qy      81 AspArgIleTyrTrpLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
Db      454 GACAGAAATCTATTACCTTCCATGCTGTGTGAGAACAGAGAAATACACTGTTTATTC 513
Qy      101 GluIleProLysThrIleAsnArgAlaValLeuMetLeuSerTrpLysProLeuLeu 120
Db      514 GAAATTTCCCAAACTATCAATAGACAGCAGCTTAATGCTCTCTTGGAGCCCTCTTTTG 573
Qy      121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg 140
Db      574 GATCTTTTTCAGGCAACACTGGCACTATGGAATGATTTCTTCAGAGAGAACTATTAAAGA 633
Qy      141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
Db      634 GAAGAAAGACCATTCGAACACTCGGAATGCTTCTTACGATTATCACCAAGGAAGTGA 693
Qy      161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
Db      694 ACATTTCTGTTTCAAGCCGGTAGTGGAAATTTATCACGTAAAGATGGAGGGCCCAAGGA 753
Qy      181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200

```

```

Db      754 TTTACGCAACAACTTTAAGGCCCAATCTAGTGGAAACTAGTTGTCCCAACATACCGGATG 813
Qy      201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
Db      814 GATCCAAAATTAATGCCCGCTGATCCAGACTGGATTGCTTTATATACATAGCAACGATATT 873
Qy      221 TrpIleSerAsnIleValThrArgGluGluArgArgLeuThrTyrValHisAsnGluLeu 240
Db      874 TGGATATCTTAACATCGTAACAGAGAGAGAGAGACTCACTTATGTGCACATGAGCTA 933
Qy      241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu 260
Db      934 GCCAACATGGAAGAGATGCCAGATCAGCTGAGTGGCTACTTGTGTTCTCCAAAGAGAA 993
Qy      261 PheAspArgTyrSerGlyTyrTrpCysProLysAlaGluThrThrProSerGlyGly 280
Db      994 TTTGATAGATATTTCTGGCTATTGGTGTCTCAAAAGCTGAACACACTCCCGTGGTGGT 1053
Qy      281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300
Db      1054 AAAATTTCTTAGAATTTCTATATGAAGAAATGATGAATCTGAGGTGGAAATATTTCATGTT 1113
Qy      301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320
Db      1114 ACATCCCTTATGTTGGAACAAAGGAGGCGAGATTCATTCCTGTTATCTTAAACAGGTACA 1173
Qy      321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
Db      1174 GCAATCTCTTAAGTCACTTTTAAGATGTGAGAAATATGATGATGCTGAAGGAGGATC 1233
Qy      341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
Db      1234 ATAGATGTCATAGATAAGGAACATAATTCACCTTTTGAGATTCTATTGGAAGGAGTTGAA 1293
Qy      361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
Db      1294 TATATTGCCAGAGCTGATGGACTCTCTGAGGAAATATGCTGTGCTCATCTACTAGAT 1353
Qy      381 ArgSerGlnThrArgLeuGlnIleValLeuLeuSerProGluLeuPheIleProValGlu 400
Db      1354 CGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATATTATTATCCCAAGTAGAA 1413
Qy      401 AspAspValMetClnArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
Db      1414 GATGATGTTATGGAAGGAGGAGACTCATTTGAGTCAGTGCCTGATCTGTGACCCCACTA 1473
Qy      421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
Db      1474 ATTATCTATGAGAAACAAACAGACATCTGGATTAATATCCATGACATCTTTCATGTTTT 1533
Qy      441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
Db      1534 CCCCAAGTCCAGAGAGGAAATTCAGTTATTTTGTGCTCTGAAATGCAAAACAGGTTTC 1593
Qy      461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
Db      1594 CGTCATTTATACAAATTTACATCTATTTTAAAGGAAAGCAAAATATAAACGATCCAGTGGT 1653
Qy      481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 500
Db      1654 GGGCTGCTCTCTCAAGTGATTTCAAGTGTCTTATCAAGAGGAGAGATAGCAATTTACCAGT 1713
Qy      501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
Db      1714 GGTGAATGGGAAGTCTTGGCCGCGATGGATCTAATATCCAAAGTTGATGAAGTCAGAAGG 1773
Qy      521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer 540
Db      1774 CTGATATATTTTGAAGGACCAAGAACTCCCTTTTAGAGCATCACCTGTAGCTAGTCACT 1833
Qy      541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560

```

Dn	1834	TACGTAATCCTGGAGAGGTGCAAGGCTGACTGACCGTGCTACTCACATTCCTTGCTGC	1893
Qy	561	IISeSrglnHisCysAspPheHeilleSerLysTyrSerAenGlnLysAenProHisCys	580
Dn	1894	ATCAGTCAGCACTGTGACTTCCTTATAAGTAAGTAGTAACACAGAGAATCCACACTGT	1953
Qy	581	ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe	600
Dn	1954	GTGTCCCTTTACAGCTATCAAGTCTGGAAGTAGACCACTTGCAAAACAAGGAATTT	2013
Qy	601	TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe	620
Dn	2014	TGGCCACCACTTTGGATTGAGTCAGCGTCTCTCTGACTACTCTCCAGAAATTTTC	2073
Qy	621	SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu	640
Dn	2074	TCCTTTGAAAGTACTACTGAGTTTACATTTGATGGATGCTCTACAGCCTCATGATCTA	2133
Qy	641	GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyProGlnValGlnLeu	660
Dn	2134	CAGCCTGGAAGAAATATCTACTGTCTGCTTCATATATGGTGGTCCCTCAGTGCAGTTG	2193
Qy	661	ValAenAenArgPheLysGlyValLysTyrPheArgLeuAenThrLeuAlaSerLeuGly	680
Dn	2194	GTGAATAATCGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCTAGCCTCTCTAGGT	2253
Qy	681	TyrValValValValIleAspAenArgGlySerCysHisArgGlyLeuLysPheGluGly	700
Dn	2254	TATGTGGTTGTAGTGATAGACAACAGGGGATCTGTCCAGGGCTTTAAATTTTGAAGGC	2313
Qy	701	AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr	720
Dn	2314	GCCTTTAATAATAAATGGGTCAAAATAGAAATGACCATCAGGTGGAAGGACTCCAAATAT	2373
Qy	721	LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyr	740
Dn	2374	CTAGCTCTCCATATGATTTTCATTTGACTTAGATCGTGTGGGCATCCAGGCTGTCTTAT	2433
Qy	741	GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle	760
Dn	2434	GGAGTATACCTCTCCGTATGGCATTAATGCAGAGTTCAGATATCTTCAGGGTGTCTATT	2493
Qy	761	AlaGlyAlaProValThrLeuTyrIlePheTyrAspThrGlyTyrThrGluArgTyrMet	780
Dn	2494	GCTGGGGCCCCAGTCACCTCTGTGGATCTTCTATGATACAGATACACGGAAGCTTATATG	2553
Qy	781	GlyHisProAspGlnAenGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu	800
Dn	2554	GGTCACCTGACCAAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAACAGAA	2613
Qy	801	LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAenVal	820
Dn	2614	AAAGTCCCCTCTGAAACCAAAATCGTTTACTGCTCTTACATGGTTTCTCGATGAGAAATGTC	2673
Qy	821	HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp	840
Dn	2674	CATTTTGCATATACCATATATTACTGAGTTTCTTTAGTAGGGCTGGAAGCCATATGAT	2733
Qy	841	LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr	860
Dn	2734	TTACAGATCTATCTCTCAGAGAGACACAGCATAGAGTTCTCGAATCGGAGAACATTAT	2793
Qy	861	GluLeuHisLeuLeuHisTyrLeuGlnGluAenLeuGlySerArgIleAlaLeuLys	880
Dn	2794	GAAGTCCATCTTTTGCACTACCTTCAGANAAACCTTGGATCAGTATTGCTGCTCTAAAA	2853
Qy	881	ValIle 882	
Dn	2854	GTGATA 2859	
Qy	AF221634	3127 bp mRNA linear PRI 05-NOV-2000	
Qy	AF221634	1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys	20
Qy	214	ATGGCAGCAGCAATGGAACAGACAGCTGGGTGTTGAGATATTGAACTGCGGACTGT	273

DEFINITION	Homo sapiens dipeptidyl peptidase 8 (DPP8) mRNA, complete cds.		
ACCESSION	AF221634		
VERSION	AF221634.1 GI:11095187		
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	1 (bases 1 to 3127) Abbott,C.A., Yu,D.M., Woollatt,E., Sutherland,G.R., McCaughan,G.W. and Gorrell,M.D. Cloning, expression and chromosomal localization of a novel human dipeptidyl peptidase (DPP) IV homolog, DPP8 Eur J Biochem. 267 (20), 6140-6150 (2000)		
AUTHORS	2 (bases 1 to 3127) Abbott,C.A., Yu,D., McCaughan,G.W. and Gorrell,M.D. Direct Submission Submitted (06-JAN-2000) A.W. Morrow Gastroenterology and Liver Centre, Centenary Institute of Cell Biology and Cancer Medicine, Locked Bag No.6, Newtown, Sydney, NSW 2042, Australia		
TITLE	Location/Qualifiers		
JOURNAL	1. 3127 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /chromosome="15" /map="15q22"		
PUBMED	1. 3127 /gene="DPP8" /gene="DPP8" 214..2862 /note="serine protease; member of prolyl oligopeptidase S9b family belonging to enzyme clan SC; alternatively spliced forms of this mRNA are found in AF221635, AF221636 and AF221637"		
AUTHORS	/codon_start=1 /product="dipeptidyl peptidase 8" /protein_id="AAC29766.1" /db_xref="GI:11095188" /translation="MAAMETEQLGWEIFETADCEINIESQDRPKLEFFYVERYSWSO LKLADTRKYHGMKAPHDPMFKRNDPDGPHSDRIYILAMSGENRENTLFSGEI PKTNRAVLMLSWKPLDLFOATLDYMSREELLRKRKRGITGVGIASDYTHQSG TFLFQAGSGYHVMDGPGQFTQQLPNLVETSCNIRMDPKLPADPDWATFIHNS DIWISNIVTREERLLTYVHNLNMEEDARSAGVATFVLOEFPDRYGVWPCFKAIT PSGKILRIIYENDESEVEI IHVTSFMLETRADSFYKPTGTANPKVTPKMSIEMI DAEGRIDVIDKELIQFEILLFEGVEYIARAGWTPGKYAWSILLDRSOTRLOIWLIS PELFTPIVEDDVMERQRLIESVPDSVPLIIEETDIIWNIHDIHFVPSHSEVIEIF IPASECTGPRHLVKITSILSKSKYKSSGGLPAPSDFKPIKEETAITSGSEVILGR HGSINVDVRLVYFEGTKDSPLEHLYVSVYVNPGEVTRLDTRGSHSCISQCHD PFISKYDNKPNCHVLYKLSPEDDPTCKTEFWATILDSAGPLPDYTPPPIFSPES TGTFTLYGMLPHDLQPKKYPTVLFIYGGQVQLVNNRFGKPYFRNLTLASLGYS VVVDNRSCHRGKLFEGAFKYGQIIEIDDOVEGLQYLASRYDFIDLDVRVIGHWSY CGYLSLMALMORSDFRVALAGAPVTLLWIFDGTGYRYMGHPDQNEQGYILGSVAMQ XKFPSEPNLLLAHGLDENVHPAHTSILLSFLVRAGKPYDLQIYQPSRHSIRVPES GSHYELHLLHYLQENLGSRTAALKVI"		
FEATURES			
source			
gene			
CDS			
Alignment Scores:			
Pred. No.:	0	Length:	3127
Score:	4700.00	Matches:	882
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	8	Gaps:	0
ORIGIN			
Alignment Scores:			
Pred. No.:	0	Length:	3127
Score:	4700.00	Matches:	882
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	8	Gaps:	0
US-10-825-632-1 (1-882) x AF221634 (1-3127)			
QY	1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20		
Db	214 ATGGCAGCAGCAATGGAACAGACAGCTGGGTGTTGAGATATTGAACTGCGGACTGT 273		

QY	21	GluGluAenIleGluSerGluAspArgProLysLeuGluProPheTyrValGluArgTyr	40	Db	1354	CGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAAATTAATTTATCCAGTAGAA	1413
Db	274	GAGGAGATATATGAATCACAGATCGGCTAAATTTGGAGCTTTTATGTTGAGCGGTAT	333	QY	401	AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu	420
QY	41	SerTyrSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet	60	Db	1414	GATGATGTTATGGAAGGACAGAGACTCATGTGAGTCAGTGCCTGATGTTCTGAGCGCCACTA	1473
Db	334	TCCTGGAGTCAGCTTAAAGAGCTCTTGGCGGATACAGAAAAATATCATGGCTACATGATG	393	QY	421	IleIleTyrGluGluThrThrAspIleTyrPheAsnIleHisAspIlePheHisValPhe	440
QY	61	AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer	80	Db	1474	ATTATCTATGAAGAAACAACAGACATCTGGATAAATATCCATGACATCTTTCATGTTTTT	1533
Db	394	GCTAAGGACCCACATGATTTTCATGTTGTGAAGAGGAATGATCCAGATGGACCTCATTTCA	453	QY	441	ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe	460
QY	81	AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer	100	Db	1534	CCCCAAGTCACGAAGAGGAAATTCAGTGTATTTTTCCTCTGCTGAATGCAAAACAGGTTTC	1593
Db	454	GACAGAACTTATACCTTGGCCATGCTCTGGTGAGAACAGAAAAATACATGTTTTATTTCT	513	QY	461	ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly	480
QY	101	GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTyrLysProLeuLeu	120	Db	1594	CGTCATTTATACAAAATTTACATCTATTTTAAAGGAAACAAATATAAACGATCCAGTGGT	1653
Db	514	GAATTTCCCAAACTATCAATAGAGAGAGCTCTTAATGCTCTCTTGGAGGCTCTTTTG	573	QY	481	GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer	500
QY	121	AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg	140	Db	1654	GGGCTCGCTGCTCCAAAGTGATTTCAAGTGCTCTATCAAGAGGAGATAGCAATTTACCAGT	1713
Db	574	GATCTTTTCAGGCAACTGACTATGGAATGATTTCTCGAGAAGAAGACTATTAAGA	633	QY	501	GlyGluTyrGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg	520
QY	141	GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly	160	Db	1714	GGTGAATGGGAAGTTCTTTGGCGGCAATGATCTAATATCCAAAGTTGATGAAGTCAGAAGG	1773
Db	634	GAAGAAGAACGATTTGAACAGCTCGGAATTCCTTCACGATTTATCACCAGAGGATGGA	693	QY	521	LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer	540
QY	161	ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly	180	Db	1774	CTGGTATATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACCTGTAGTAGTACGT	1833
Db	694	ACATTTCTGTTTCAAGCCGCTAGTGGAAATTTATCACGTAAGATGAGGAGGCGCACAGA	753	QY	541	TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys	560
QY	181	PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet	200	Db	1834	TAGTAAATCTCTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACATTTCTTGCTGC	1893
Db	754	TTTACGCCAACACCTTTAAGGCCCAATCTAGTGGAAATCTGTTTATACATAGCAACGATAT	813	QY	561	IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys	580
QY	201	AspProLysLeuCysProAlaAspProAspTyrPheAlaPheIleHisSerAsnAspIle	220	Db	1894	ATCAGTCAGCACTGTGACTTCTTTTATAAGTAAGTATAGTACCAGAGAAGATCCACACTGT	1953
Db	814	GATCCAAAATATATGCCCGCTGATCCAGACTGGATGCTTTTATACATAGCAACGATAT	873	QY	581	ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe	600
QY	221	TyrIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu	240	Db	1954	GTGTCCTCTTACAAAGCTATCAAGTCTCTGAAGATGACCAACTTCGCAAAACAAGGAATTT	2013
Db	874	TGGATATCTTAACATCGTAACACAGAGAGAGAGAGACTCACTTATGTGCACATGAGCTA	933	QY	601	TyrAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe	620
QY	241	AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu	260	Db	2014	TGGGCCACCAATTTTGGATTCAGCAGGCTCTCTCTGACTATATCTCTCCAGAAATTTTC	2073
Db	934	GCCAACTGGAAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAGAGAA	993	QY	621	SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu	640
QY	261	PheAspArgTyrSerGlyTyrTyrTyrPyrCysProLysAlaGluThrThrProSerGlyGly	280	Db	2074	TCCTTTTGAAGTACTACTGGATTTTACATTTGTATGGATGCTCTCAAGCCTCATGACTA	2133
Db	994	TTTGATAGATATCTGGCTATTTGGTGTGTCCTCAAGCTGAACCACTCCAGTGGTGT	1053	QY	641	GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu	660
QY	281	LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHisVal	300	Db	2134	CAGCTCGGAAGAAATATCTACTGCTGTGTTTATATATGTTGTTGCTCAGTGCAGTTG	2193
Db	1054	AAAACTTTAGAAATCTATATGAAGAAATGATCAATCTGAGTGGAAATATTATCATGTT	1113	QY	661	ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly	680
QY	301	ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr	320	Db	2194	GTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGTTTGAATACCTAGCCTCTCTAGT	2253
Db	1114	ACATCCCTTATGTTGAAACAAGGAGGCGAGATTCATTCGCTTATCTCAAAACAGGTACA	1173	QY	681	TyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly	700
QY	321	AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle	340	Db	2254	TATGTGTTGTAGTGATAGACAAACAGGGGATCTGTCTACCCAGGCGCTTAAATTTGAAGGC	2313
Db	1174	GCAAAATCTTAAAGTCACTTTTAAAGATGTCAGAAATAATGATTTGATGCTGAAGAGGATC	1233	QY	701	AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr	720
QY	341	IleAspValIleAspLysGluIleGlnProPheGluIleLeuPheGluGlyValGlu	360	Db	2314	GCCTTTAAATATATAAATGGGTCAAATAGAAATTCAGATAGGTCAGGTCAGGCTCCAATAT	2373
Db	1234	ATAGATGTCATAGATAAGGAACTAATCAACTTTTGGATTCATTTTGAAGAGGTTGAA	1293	QY	721	LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyr	740
QY	361	TyrIleAlaArgAlaGlyTyrThrProGluGlyLysTyrAlaTyrSerIleLeuLeuAsp	380	Db	2374	CTAGCTTCGATATGATTTTCACTTAGATGCTGTGGGATCCAGGCTGGTCTTAT	2433
Db	1294	TATATTCAGAGAGCTGGATGAGCTCTGAGGGAATAATGCTTGGTCCATCTACTAGAT	1353	QY	741	GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle	760
QY	381	ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu	400				

Db	2434	GGAGGATACCTCTCCCTGATGGCAATTAATGCAGAGGTGAGATATCTTCAGGGTTCCTATT	2493
Qy	761	AlaGlyAlaProValThrLeuThrPheTyrAspThrGlyTyrThrGluArgTyrMet	780
Db	2494	GCTGGGGCCCGAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACGTTATATG	2553
Qy	781	GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu	800
Db	2554	GGTCACCTGACCAAGATGAACAGGGCTATTACTTATAGGATCTGTGGCCATGCAACGAGAA	2613
Qy	801	LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal	820
Db	2614	AAGTCCCTCTGAACCAAAATCGTTACTGCTCTTACATGGTTTCTCGATGAGAATGTC	2673
Qy	821	HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp	840
Db	2674	CATTTGACATACACAGTATATTACTGAGTTTTTTAGTGAGGCTGGAAGCCATATGAT	2733
Qy	841	LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr	860
Db	2734	TTACAGATCTATCTCTAGGAGACACAGCATTAAGAGTTCTCTGNAATCGGAGACATTTAT	2793
Qy	861	GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLys	880
Db	2794	GAACGTGATCTCTTTGCACTACCTTCAAGAAAAACCTTGGATCAGTATTGCTCTTAAAA	2853
Qy	881	Vallie	882
Db	2854	GTGATA	2859
RESULT 9			
AX354793		3143 bp	DNA linear PAT 06-FEB-2002
LOCUS		Sequence 1 from Patent WO0179473.	
DEFINITION		AX354793	
ACCESSION		AX354793.1	GI:18619526
VERSION			
KEYWORDS		Homo sapiens (human)	
SOURCE		Homo sapiens	
ORGANISM		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominoidea; Homo.	
REFERENCE		1	Meyers, R.A. and Williamson, M.
AUTHORS		21953, a human prolyl oligopeptidase family member and uses thereof	
TITLE		Patent: WO 0179473-A 1 25-OCT-2001.	
JOURNAL		Millennium Pharmaceuticals, Inc. (US)	
FEATURES		Location/Qualifiers	
source		1..3143	
		/organism="Homo sapiens"	
		/mol_type="unassigned DNA"	
		/db_xref="taxon:9606"	
		229..2877	
		/note="unnamed protein product"	
		/codon_start=1	
		/protein_id="CAD22990.1"	
		/db_xref="GI:18619527"	
CDS		/translation="MAAAMETEQLGVIEIFETACDENIESQDRPKLEPFYVERYSNSQ LKKLLADTRKYHYGMAMAKPHDFMVKRNDPDGPHSDRIYILAMSGENRENTLFYSEI PKTINRAVLMSLWKPLDLFOATLDYGMYSREELLRERKRIGTVGASVDYHGSG TFLFOAGSGIYHVVDGPGQFTQOPLRNVLVETSCFNIRMDPKLCPADPDWTFATHSN DIWISNIVREERLTYVHNELANMEEDARSAGVATFVLQEEFDYRGVYWCPCAEET PSGGKILALYEENDESEVEIHTVSMLETRADSFYKPTGTANPKVTFKMSIEMI DAEGRIIDVIDKEILQPIELFEGVEYIARAGWTPEGKAWMSILLDRSQRLQIVLIS PELFTPVDDWQERQRLTISVPDSPTLIYEETDWINIHDIFHFVQSEHEEIEF IFASCKTGFRHLVYKITSILKESKYKRSGLPAPSDFKPIKEIATITSGEWEVLGR HGSNQVDEPHLVYVFEQTKDSPLEHLVYVSVNFEVTRLTDRGYSHSCCISQHCD TFISKYGNKPHPCVSLYKLSPPEDPTCKEPMATILDSAGPLDPYTPPEISFSES TTGFTLYGMLYKPHLOPKKYPTVLFYGGPOVLNNRFKGVYFRNLNTLASLGVY VVVIDNCSHRGKLFKFGAPKYMGQIEIDDOVGLQYLASVDFDLDLDRVIGHWSY GVGLSLMALMORSDI PRVALAPVTLWIFDYDTGYTRYMGHPDQNEQCYVLGVSAMQ AKKFPSEPNLLALHGLDENVHPAHTSILLSFLVRAGKPYDLQIYVQERHSRVPE GEHYELHLLHYLQENLGRINALVKI"	

Db 1129 ACATCCCTATGTTGGAAACAGAGGAGGAGATTCATTCGCTTATCCCTAAACAGGTACA 1188
Qy 321 AlaAsnProLysValThrPheLysMetSerGluLeuMetIleAspAlaGluGlyValGile 340
Db 1189 GCAATCTCTAAAGTCACTTTTAAGATGTCAGAAATATGATGATGCTGAGGAGGATC 1248
Qy 341 IleAspValIleAspLysGluLeuIleGlnProPheGluLeuPheGluGlyValGlu 360
Db 1249 ATAGATGTCATAGATAGGAACATAATCAACCTTTTGAATCTTATTTGAAGAGGTTGAA 1308
Qy 361 TyrIleAlaArgAlaGlyThrThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
Db 1309 TATATTCAGAGCTGGATGGACTCCTCGAGGAGAAATATGCTTGGTGCATCTCTAGAT 1368
Qy 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
Db 1369 CGCTCCAGACTCCCTCGAGATAGTGTGATCTCACCTGATTAATTTATCCAGTAGAA 1428
Qy 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
Db 1429 GATGATGTTATGGAAGAGCAGAGACTCATTGAGTCAGTGCCTGATTTCTGTGAGGCCACTA 1488
Qy 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleIleAspIlePheHisValPhe 440
Db 1489 ATTATCTATGAAGAAACACAGACATCTGGATTAATATCCATGACATCTTTTCATGTTTT 1548
Qy 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
Db 1549 CCCCAAGTCAAGAGAGGAAATGAGTTTATTTTGGCTTCTGATGCAAAACAGGTTTC 1608
Qy 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
Db 1609 CGTCAATTAACAATAATACATCTATTTTAAAGGAAGCAATATAAAGATCCAGTGGT 1668
Qy 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluLeuAlaIleThrSer 500
Db 1669 GGGCTGCTGCTCCAAAGTATTTCAAGTGTCTTCAAGAGAGGAGATAGCAATATACCAGT 1728
Qy 501 GlyGluTrpGluValLeuLysArgHisGlySerAsnIleGlnValAspGluValArg 520
Db 1729 GGTGAATGGGAAGTCTTGGCCGCGCATGGATCTAAATATCAAGTTGATGAAGTCAGAAG 1788
Qy 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSer 540
Db 1789 CTGGTATATTTTGAAGGACCAAAAGATCCCTTTAGAGCATCACCTGTACGTAGTCAGT 1848
Qy 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
Db 1849 TAGCTAAATCTCGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACATCTTGTCTGC 1908
Qy 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
Db 1909 ATCAGTCAGCAGCTGTGACTCTCTTTATAGTAAAGTATAGTAACCAAGAAATCCACACTGT 1968
Qy 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
Db 1969 GTGTCCCTTTACAGCTATCAAGTCCCTGAGATGACCACTTGGCAAAACAAAGGATTT 2028
Qy 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluLeuPhe 620
Db 2029 TGGGCCACCAATTTGGATTCAGCAGGCTCTCTCTGACTATACTCTCCAGAAATTTTC 2088
Qy 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
Db 2089 TCTTTTGAAGTACTACTGCAATTTACATTTGATGGATGCTCTACAGGCTCATGATCTA 2148
Qy 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyLysProGlnValGlnLeu 660
Db 2149 CAGCCTGGAAGAAATATCTTACTGTGCTGTTCATATATGGTGGTCTCAGGTCAGGTTG 2208
Qy 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680

Db 2209 GTGAATAATCGGTTAAAGGAGTCAAGTATTTTCGGCTTGAATACCTAGCCTCTCTAGGT 2268
Qy 681 TyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 700
Db 2269 TATGTGGTGTAGTGCATAGACAACAGGGGATCCTGTCCAGGAGGCTTAAATTTGAAGGC 2328
Qy 701 AlaPheLysTyrLysMetGlyGlnIleGluLeuAspGlnValGluGlyLeuGlnTyr 720
Db 2329 GCCTTTAAATATAAATGGTCAAAATAGAATTTGACCATCAGGTGGAGGACTCCATAT 2388
Qy 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyr 740
Db 2389 CTAGCTTCTCGATATGATTTTCATTTAGCTAGATCGTGTGGCATCCACGGCTGTCTAT 2448
Qy 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
Db 2449 CGAGGATACCTCTCCCTGATGGCATTAATGCAGAGGTTCAGATATCTTTCAGGGTGTCTATT 2508
Qy 761 AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780
Db 2509 GCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGATACACGGACGTTATATG 2568
Qy 781 GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
Db 2569 GGTCACTCCCTGACAGATGAACAGGCTATTTACTTAGGATCTGTGGCCTATGCAAGCAGAA 2628
Qy 801 LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
Db 2629 AAGTTCCCTCTGAACCAATCGTTACTGTCTTACATGGTTTTCTTGGATGAGAATGTC 2688
Qy 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
Db 2689 CATTTTGCACATACAGATATATTTACTGAGTTTTTTAGTGGGCTGTGAAGCCATATGAT 2748
Qy 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
Db 2749 TTACAGATCTATCTCAGAGAGACACAGCATAGAGTCTCTGAATCGGAGAACATTA 2808
Qy 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeuLys 880
Db 2809 GAAGTCATCTTTTGCATCTACCTTCAGAAAACCTTGGATCAGCTATTTGCTGCTCTAAA 2868
Qy 881 ValIle 882
Db 2869 GTGATA 2874

RESULT 10

BC040203
LOCUS
DEFINITION
IMAGE:4824813, complete cds.
ACCESSION
BC040203
VERSION
BC040203.1
KEYWORDS
MGC.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 4353)
REFERENCE
AUTHORS
Srausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shennan, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heieh, F.,
Diatchenko, L., Marusina, K., Farmer, A., Rubin, G.M., Hong, L.,
Scapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Tosnyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.Y.A., Peters, G.J.,
Abrahamson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,

BC040203 4535 bp mRNA linear PRI 07-OCT-2003
Homo sapiens dipeptidylpeptidase 8, mRNA (cdna clone MGC:42570
IMAGE:4824813), complete cds.

Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Bartels, Y.S., Kozminski, M.I., Skalek, U., Smal, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2. (bases 1 to 4535)
Strausberg, R.
Direct Submission
Submitted (27-NOV-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs-x@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunsberg, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulsegod, H.,
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
A.N., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAC Plate: 71 Row: k Column: 23.
Location/Qualifiers
1. .4535
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:42570 IMAGE:4824813"
/tissue_type="Testis"
/clone_lib="NIH MGC 97"
/lab_host="DH10B"
/note="Vector: pBluescript"
1. .4535
/gene="DPP8"
/note="synonyms: FLJ20283, FLJ14920, MGC26191"
/db_xref="GeneID:54878"
/db_xref="MIM:606819"
1629..4277
/gene="DPP8"
/codon_start=1
/product="DPP8 protein"
/protein_id="AAH40203.1"
/db_xref="GI:26007917"
/db_xref="GeneID:54878"
/db_xref="MIM:606819"
/translations="MAAAMETSQLGVEIFETADCEENIESQDRPKLPFFVRYSWSQ
LKLLADTRYHGMKAPDFMFKVNRDPDGHSDRIYLYLAMSNGENRNTILFYSEI
PKTINRAVLMLSWKLLDLFOALDYGMYSREELRKRIGTVGLASYDYHQSGD
TFIFQAGSVGHVKGQPGQTQPLRPNLVETSCFNIRMDPKLCPADPDWPIHGN
DIWISNIPREERLYVHNELANMEEDARSAGVAIFVLQEEFDRIYSGYWCPKAETI
PSGKILRLYEENDESEVEIHHVTSPLMETRADSPRYKGTANPKVTFKMEIMI
DAGRIIDVLEKLIQFFELFEGVEYARAGWTEPGYAWSIILDRSQRLQVILIS
PELFIPEVDDEMRQRLIESVPSVTPLIYBETTDIWINTHIDIFHVPQSHHEEVLIR
IFASECTGRPHLYKITSILKESKYKRSGLPAPSPDKPCIKBEIAITSGEEVLGR
HGNIQVDEVRNVYFEGTKDSPLEHLHYVSVYVNPGEVRLTDRGYSHSCISQCHD
FFYSKYNQKNRCHVLSYGLKSSPDDPTCKTFKEFWATILDSAGLPDPVTPPEIFSFES
TTGFTLYGMLYKPHDLQPGKQYVLFIFYGGPVQVNNRPFKGVKLYFRNLTLASLVY
VVINDRGSCHRLGKFGKAFKYKMGQIEIDQVQGLQVLAISRYDFILDRVGIHMSY
GGYLSMALMQRSDIIRVAIAGPVTWTFYDTGYTERYMGHPQNEQGYVLSVAMQ
AEKFPSPNRLLLHGFLDENVHFAHTSILLSFLVRAGKPYDLQIYPOERHSIRVPS

misc_feature		GEHYELHLHYLQENLGSRIAALKVI" 3207..4196 /gene="DPP8" /note="DAP2: Region: Dipeptidyl aminopeptidases/acylaminoacyl-peptidases [Amino acid transport and metabolism]" /db_xref="CDD:COG1506"	
ORIGIN			
Alignment Scores:		Length: 4535	
Pred. No.:		Matches: 881	
Score:		4696.00	
Percent Similarity:		99.9%	
Best Local Similarity:		99.9%	
Query Match:		Indels: 0	
DB:		Gaps: 0	
US-10-825-632-1 (1-882) x BC040203 (1-4535)			
QY	1	MetAlaAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys	20
Db	1629	ATGGCAGCAGCAATGGAAACAGAACACAGCTGGGTGTGGAGATATTTGAAACTCGGACTGT	1688
QY	21	GluGluAenIleGluSerGlnAspArgProLyLeuGluProPheTyrValGluArgTyr	40
Db	1689	GAGGAGAAATTTGAATCACAGGATCGGCTTAAATTCGAGCCTTTTATGTGTGAGCGGTAT	1748
QY	41	SerTTPSerGlnLeuLysLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet	60
Db	1749	TCTTGGAGTCAGCTTAAAGAGCTGCTTGGCAGTACACAGAAATATCATCGCTACATGATG	1808
QY	61	AlaIleValProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer	80
Db	1809	GCTAAGGACCAACATGATTTTCATGTTTGTGAAAGAAATGATCCAGATGAGCTCATTC	1868
QY	81	AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer	100
Db	1869	GACAGATCTAATACCTTGGCATGCTGGTGAGAGACAGAGAAATACATGTTTATTTCT	1928
QY	101	GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu	120
Db	1929	GAAATTTCCAAAACTATCAATAGACGACAGCTCTTAATGCTCTCTTTGGAAGGCTCTTTTG	1988
QY	121	AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg	140
Db	1989	GATCTTTTTCAGGCACACTGGACTATGGAATGTATTTCTCGAAGAGAAAGAACTATTAA	2048
QY	141	GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly	160
Db	2049	GAAAGAAAACGCAATGGACAGTCGGAATTTGCTTTACGATTTATCACCAGGAAGTGGG	2108
QY	161	ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly	180
Db	2109	ACATTTCTGTTTCAAGCGGTAGTGGAAATTTATCAGTAAAGATGGAGGCCACAAAG	2168
QY	181	PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet	200
Db	2169	TTTACGCAACCACTTTTAGCCCACTAGTGGAAACTAGTTGTCCTCCCACTACGATG	2228
QY	201	AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle	220
Db	2229	GATCCAAATATTCCTCTGCTGATCCAGCTGGATGCTTTATACATGGCAACGATAT	2288
QY	221	TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisGlnLeuLeu	240
Db	2289	TGGATATCTAACTCGTAACAGAGAAAGAGAGACTCCTTATGTGCACAAATAGCTA	2348
QY	241	AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu	260
Db	2349	GCCAAACATGGAAGAGATGCCAGATCAGCTGGAGTGCCTTCTTCTTCCAAAGAA	2408
QY	261	PheAspArgTyrSerGlyTyrTyrTyrTyrTyrTyrTyrTyrTyrTyrTyrTyrTyr	280

Db 2409 TTTGATAGATATCTGGGCTATTGGTGGTGTGTCCTCAAAAGCTGAAACAATCTCCAGTGGTGT 2468
Qy 281 LysileLeuArgileLeuTyrGluGluAsnAspGluSerGluValGluIleleHisVal 300
Db 2469 AAAATCTTAGAATTTCTATATGAAGAAATGATGATCTGAGTGGAAATATTCTATGTT 2528
Qy 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320
Db 2529 ACATCCCTATGTTGGAAACAAGGAGGCGAGATTCATTCGGTTATCCTAAACAGGTACA 2588
Qy 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
Db 2589 GCAATCTCTAAAGTCACTTTTAAGATGTGAGAAATATGATGCTGCTGAAGAGATC 2648
Qy 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
Db 2649 ATAGATGTCATAGATAGAGAACTAAATTCACCTTTTGAGATCTTATTTGAAGAGATTGA 2708
Qy 361 TyrIleAlaArgAlaGlyTyrThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
Db 2709 TATATGGCAGAGCTGGATGACTCTCTGAGGAAATATGCTGGTCCATCTCTACTAGAT 2768
Qy 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
Db 2769 CGCTCCAGACTCGCTACAGATAGTGTGATCTCACTGCAATATTATTCACAGTAGAA 2828
Qy 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
Db 2829 GATGATGTTATGGAAAGCAGAGACTCAATTGAGTCAGTGGCTGCTGTCGAGCCACTA 2888
Qy 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
Db 2889 ATTATCTATGAAGAAACAACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTT 2948
Qy 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
Db 2949 CCCCAAGTCAAGAAAGAGAAATTCAGTTTATTTTTCCTCTGAATGCAAAACAGGTTTC 3008
Qy 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
Db 3009 CGTCAATTTATACAAATTTACATCTATTTTAAAGAAAGCAAAATATAACGATCCAGTGT 3068
Qy 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluIleAlaIleThrSer 500
Db 3069 GGGCTGCCTCCCAAGTGATTTCAAGTGCTCTATCAAGAGAGATAGCAATTCACAGT 3128
Qy 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
Db 3129 GGTGAATGGGAAGTTCTTGGCGGCGATGGATCTAATATCCAGTTGATGAGTCAGAGG 3188
Qy 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLysLeuTyrValValSer 540
Db 3189 CTGGTATATTTTGAAGGACCCAAAGACTCCCTTTAGAGCATCACCTGTACGTAGTCACT 3248
Qy 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
Db 3249 TAGGTAATCTCTGGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACATTTCTGCTGC 3308
Qy 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
Db 3309 ATCAGTCAGCACTGTGACTTCTTTAAGTAAAGTATAGTAACCCAGAGAAATCCACACTGT 3368
Qy 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
Db 3369 GTGTCCCTTTACAGCTATCAAGTCTCAAGATGACCCCACTTGCACAAACAAAGGAATTT 3428
Qy 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 620
Db 3429 TGGGCCACCAATTTGGATTCAGAGGTCCTCTTCTGACTATATCTACTCTCCAGAAATTTTC 3488
Qy 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
Db 3489 TCTTTTGAAGTACTACTGGATTTACATTTGATGGGATGCTCTACAAGCTCTCATGATCTA 3548

Qy 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 660
Db 3549 CAGCCTGGAAAGAAATATCTACTGTGTTCTATATGATGGTGTCTCAGGTCACTTG 3608
Qy 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680
Db 3609 GTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGTTGAATACCTTAGCCTCTCTAGGT 3668
Qy 681 TyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 700
Db 3669 TATGTGGTGTGATGATAGACCAACAGGGGATCCTGTCAACCAGGGCTTTAAATTTGAAGGC 3728
Qy 701 AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr 720
Db 3729 GCCTTTAAATATAAATGGGTCAATAGAAATGACGATCAGGTGGAGAGACTCCATAT 3788
Qy 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyr 740
Db 3789 CTAGCTTCTCGATATGATTTTCACTTAGATCGTGGGCATCCACGGCTGCTCTAT 3848
Qy 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
Db 3849 GGAGGATACCTCTCCCTGATGGCAATTAATGCAGAGGTCAAGATATCTTCAGGGTTGCTATT 3908
Qy 761 AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780
Db 3909 GCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACGTTATATG 3968
Qy 781 GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
Db 3969 GGTCACTCCGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATCGAAGCGAA 4028
Qy 801 LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
Db 4029 AAGTTCCTCTGGAACCAATCGTTTACTGCTCTTACATGGTTTCTTGGATGAGAAATGTC 4088
Qy 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
Db 4089 CATTTTGCATACACCATATATTTACTAGTTTTTTAGTGGGCTGGAAAGCCATATGAT 4148
Qy 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
Db 4149 TTACAGATCTATCTCTCAGGAGACACAGCATAGAGTTCTCTGAATCGGAGAACATTAT 4208
Qy 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLys 880
Db 4209 GAACCTGATCTTTTGCACTTACCTTCAAGAAACCTTGGATCAGCTATTGCTGCTCTAAAA 4268
Qy 881 ValIle 882
Db 4269 GTGATA 4274

RESULT 11

AR631283 LOCUS 4829 bp DNA linear PAT 14-FEB-2005
DEFINITION Sequence 12 from patent US 6844180.
ACCESSION AR631283
VERSION AR631283.1 GI:59770927
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 4829)
AUTHORS Oi, S., Akinsanya, K.O., Riviere, P.J.M. and Junien, J.-L.
TITLE Serine protease genes related to DPPIV
JOURNAL Patent: US 6844180-A 12 18-JAN-2005;
Ferring BV;
NLM; Location/Qualifiers
source 1. 4829
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN	Db	1114	ACATCCCTATGTTGGAAACAAGGAGGCGAGATTCTCCGTTATCTCTAAAAAGGTACA	1173
Alignment Scores:	Qy	321	AlaAsnProLysValThrPheLysMetSerGluLeuMetIleAspAlaGluGlyArgIle	340
Pred. No.:	Db	1174	GCAAAATCCTAAAGTCACCTTTTAAAGATGTCAGAAATATGATGATGCTCAGAGGAAGATC	1233
Score:	Qy	341	IleAspValIleAspLysGluLeuGlnProPheGluIleLeuPheGluGlyValGlu	360
Percent Similarity:	Db	1234	ATAGATGTCATAGATAAGGAACATTAATCAACCTTTTGAATCTATTTAGAGAGTTGAA	1293
Best Local Similarity:	Qy	361	TyrIleAlaArgAlaGlyTyrThrProGluGlyLysTyrAlaThrPheSerIleLeuLeuAsp	380
Query Match:	Db	1294	TATATTCGACAGCTGGATGGACTCTCGAGGGAATAATGCTTGGTCCATCTCTACTAGAT	1353
DB:	Qy	381	ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu	400
US-10-825-632-1 (1-882) x AR631283 (1-4829)	Db	1354	CGCTCCAGACTCGCCTACAGATAGTGTGATCTCACTCCTGAATATTTATATCCAGTAGAA	1413
Qy	Qy	401	AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu	420
Qy	Db	1414	GATGATGTTATGGAAGGACAGAGACTCATTTGAGTTCAGTGCCTGATTTCTGTGACGCCACTA	1473
Qy	Qy	421	IleIleTyrGluThrThrAspIleTyrPheAsnIleHisAspIlePheHisValPhe	440
Qy	Db	1474	ATTATCTATGAAGAAACAACAGACATCTGGATAAATATCCATGCATCTTTCTGATGTTT	1533
Qy	Qy	441	ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe	460
Qy	Db	1534	CCCCAAGTCACGAAGAGGAATTTGAGTTATTTTGGCTCTGAATGCAAAACAGGTTTC	1593
Qy	Qy	461	ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysValArgSerSerGly	480
Qy	Db	1594	CGTCATTTATACAAATTTACATCTATTTTAAAGGAAACAAATATAAACGATCCAGTGGT	1653
Qy	Qy	481	GlyLeuProAlaProSerAspPheLysCysProLysGluGluIleAlaIleThrSer	500
Qy	Db	1654	GGCTGCCTGCCCAAGTATTTCAAGTGTCTTCAAGAGGAGATAGATGATTTACAGT	1713
Qy	Qy	501	GlyGluTyrGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg	520
Qy	Db	1714	GGTGAATGGGAAGTTCTTTGGCCGCGATGATCTAATATCCAAAGTTGAAGTCAAGAGG	1773
Qy	Qy	521	LeuValThrPheGluGlyThrLysAspSerProLeuGluHisLysLeuTyrValSer	540
Qy	Db	1774	CTGGTATATTTTGAAGGCGACCAAGACTCCCTTTTAGAGCATCACCTGTACGTAGTCAGT	1833
Qy	Qy	541	TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys	560
Qy	Db	1834	TACGTAAATCCCTGGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACATTTCTTGCTGC	1893
Qy	Qy	561	IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys	580
Qy	Db	1894	ATCAGTCAGACTGTGACTTCTTTTATAAGTAAGTATAGTAACCAAGAAATCCACACTGT	1953
Qy	Qy	581	ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe	600
Qy	Db	1954	GTGTCTTTTACAGCTATCAAGTCTGAGATGACCCCACTTCGAAACCAAGGAATTT	2013
Qy	Qy	601	TyrAlaThrIleLeuAspSerLysGluProLeuProAspTyrThrProGluIlePhe	620
Qy	Db	2014	TGGGCCACCATTTTGAATTCAGAGTCTCTTTCTGACTATATCTCTCCAGAAATTTTC	2073
Qy	Qy	621	SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu	640
Qy	Db	2074	TCTTTTGAAGATCTACTGATTTTACATTTGATGGATGCTCTCAAGCCCTCATGATCTA	2133
Qy	Qy	641	GlnProGlyLysLysTyrProThrValLeuPheIleTyrGly--ProGlnValGlnL	660
Qy	Db	2134	CAGCTCGGAAGAAATATCTACTGTGCTGTTTCATATATGTTGGTCTCTCAGGTGAGT	2193
Qy	Qy	660	euValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrIleuAlaSerLeuG	680

Db 1174 GCAAACTCTAAAGTCACTTTTAAAGATGTGAGAAATAATGATGTGCTGAGGAAGGATC 1233
 Qy 341 ileAspValIleAspIleAspIleGluIleGluIleGluIleLeuPheGluIleValGlu 360
 Db 1234 ATAGATGTCATAGTAAGGAACCTAAATTCACCTTTTGGAGATCTTATTTGAAGGAGTTGAA 1293
 Qy 361 TyrIleAlaArgAlaGlyTTPThrProGluGlyIleValAspIleLeuLeuAsp 380
 Db 1294 TATATTGCCAGAGCTGGATGCTCTGAGGGGAAATAATGCTTGGTCCATCTACTAGAT 1353
 Qy 381 ArgSerGlnThrArgLeuGluIleValLeuIleSerProGluLeuPheIleProValGlu 400
 Db 1354 GCCTCCAGACTCCCTACAGATAGTGTGATCTCACCTGATTTATTTATCCAGTAGAA 1413
 Qy 401 AspAspValMetGluArgGluIleGluIleGluIleGluIleGluIleGluIleGluIle 420
 Db 1414 GATGATGTTATGGAAAGGAGAGACTCATTTGATCTGCTGATCTGCTGATCTGCTGATCT 1473
 Qy 421 IleIleTyrGluGluThrThrAspIleTTPIleAsnIleHisAspIlePheHisValPhe 440
 Db 1474 ATTATCTATGAAGAAACAAAGACATCTGATATAATATCCATGACATCTTTTCATGTTT 1533
 Qy 441 ProGlnSerHisGluGluIleGluIleGluIlePheIleAspIleGluIleGluIleGluIle 460
 Db 1534 CCCCAAGTCAAGAGAGAAATGAGTTTATTTTGGCTCTGAAATGCAAAACAGGTTTC 1593
 Qy 461 ArgHisLeuTyrIleThrSerIleLeuLeuGluSerIleTyrIleArgSerSerGly 480
 Db 1594 CGTCATTTATACAAATATACATCTATTTTAAAGAGAAACAAATAATAACGATCCAGTGGT 1653
 Qy 481 GlyLeuProAlaProSerAspPheIleGluIleGluIleGluIleGluIleGluIleGluIle 500
 Db 1654 GGGCTCGCTCCCAAGTCAATTTTCAAGTGTCTTATCAAGAGAGAGATGCAATATACCAAGT 1713
 Qy 501 GlyGluTTPGluValLeuGlyIleArgHisGlySerAsnIleGluIleValAspGluValArg 520
 Db 1714 GGTGAATGGAGGATCTTTGGCGGCATGGATCTAATATCCAAGTTGATCAAGTCAAGAAG 1773
 Qy 521 LeuValTyrPheGluIleGluIleGluIleGluIleGluIleGluIleGluIleGluIleGluIle 540
 Db 1774 CTGGTATATTTGAAGGACCAAGAGACTCCCTTTAGAGCATCACCTGTACGTAGTCAAGT 1833
 Qy 541 TyrValAsnProGluIleGluIleGluIleGluIleGluIleGluIleGluIleGluIleGluIle 560
 Db 1834 TAGGTAATCTCTGGAGAGTGAACAGGCTGACTGACCGTGGCTACTCACATTTCTTGTCTG 1893
 Qy 561 IleSerGluHisCysAspPhePheIleSerIleTyrSerAsnGlnIleValAsnProHisCys 580
 Db 1894 ATCAGTCAGACTGTGACTCTTTATAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTA 1953
 Qy 581 ValSerLeuTyrIleLeuSerSerProGluAspAspProThrCysIleThrIleGluPhe 600
 Db 1954 GTGTCCCTTTACAAGCTATCAAGTCTGAAGATGACCACTTGCATGCAAGCAAGCAAGCAAG 2013
 Qy 601 TrpAlaThrIleLeuAspSetAlaGlyProLeuProAspTyrThrProGluIlePhe 620
 Db 2014 TGGGCCACCATTTTGGATTCAGCAGGCTCTCTCTGACATATCTCTCAGAAATTTTC 2073
 Qy 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrIleProHisAspLeu 640
 Db 2074 TCTTTTGAAGTACTACTGGATTTACATTTGATGGATGCTCTCAAGCCTCATGATCTA 2133
 Qy 641 GlnProGlyIleValIleTyrProThrValLeuPheIleTyrGlyGly--ProGlnValGlnIle 660
 Db 2134 CAGCCTGGAAAGAAATATCTACTGTCTGTGTTTCATATATGTTGCTCTCTCAGGTGACGT 2193
 Qy 660 euValAsnAsnArgPheIleGlyValIleValIleTyrPheArgLeuAsnThrLeuAlaSerLeu 680
 Db 2194 TGGTGAATATCGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTAGCCTCTCTCTAG 2253
 Qy 680 lYrTyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLeuPheGluG 700
 Db 2254 GTTATGTGGTTGTAGTATAGACAAACAGGGGATCTCTGTCAACCGAGGGCTTAAATTTGAAG 2313

Qy 700 lYAlaPheIleTyrIleMetGlyGlnIleGluIleLeuAspGlnValGluGlyLeuGlnT 720
 Db 2314 CGGCCTTTAATATAAATGCGGTCAATAGAAATTTGACGATCAGGTGGAAGGACTCCAAT 2373
 Qy 720 YrLeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleIleGlyTTPSerT 740
 Db 2374 ATCTAGCTTCTCGATATGATTTTCATTTGACTTATAGATCGTGTGGCATCCACGGCTGGCTCT 2433
 Qy 740 YrGlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaI 760
 Db 2434 ATGAGAGATACCTCTCCCTGATGGCATTAATCAGAGGTGAGATATCTTCAGGGTTGCTA 2493
 Qy 760 leAlaGlyIleValProValThrLeuTTPIlePheTyrAspThrGlyTyrThrGluArgTyrM 780
 Db 2494 TTGCTGGGGGCCCGCCAGTCACTCTGTGGATCTTCTATGATACAGATACAGGACGTATATA 2553
 Qy 780 etGlyIleHisProAspGlnAsnGluIleGlyTyrTyrIleGlySerValAlaMetGlnAlaG 800
 Db 2554 TGGGTCACTCCCTGACCAAGATGAACAGGGCTATTACTTAGGATCTCTGGCCATGCAAGCAG 2613
 Qy 800 lYlYsPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnV 820
 Db 2614 AAAAGTTCCCTCTGAAACCAATCGTTTACTCTTACATGCTTCTCTGATGAGAATG 2673
 Qy 820 alHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyIleYsProTyrA 840
 Db 2674 TCCATTTTGACATACACCACTATATTACTGATGTTTTTACTGAGGGCTGGAAGCCATATG 2733
 Qy 840 pLeuGlnIleTyrProGlnGluArgHiserIleArgValProGluSerGlyIleGluHisT 860
 Db 2734 ATTACAGATCTATCTCTCAGAGAGACACAGCATAAGAGTTCTCTGAATCGGAGAACAT 2793
 Qy 860 YrGluLeuHisLeuLeuLeuHisTyrIleGlnGluAsnLeuGlySerArgIleAlaAlaLeuL 880
 Db 2794 ATGAATCTCATCTTTTGACATCTACTCTCAAGAAACCTTGGATCAACGATTTGTCTCTAA 2853
 Qy 880 ySValIle 882
 Db 2854 AAGTGATA 2861

RESULT 13
 BC043124
 LOCUS
 DEFINITION Mus musculus dipeptidylpeptidase 8, mRNA (cdna clone MGC:58057
 IMAGE:6410075), complete cds.
 BC043124
 VERSION BC043124.1 GI:27695449
 KEYWORDS MGC.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Murioidea; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 4799)
 AUTHORS Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Frange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Meny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez J., Helton E., Madan A., Young A.C., Rodriguez S., Paney A., Whittington M., Madan A., Young A.C., Shevchenko Y., Bouffard G.O., Blakeley R.W., Touchman J.W., Green E.D., Buckson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S., Krzyzanski M.I., Skalek U., Smaluk D.E., Schnerch A., Schein J.E., Jones S.J. and Marra M.A.

TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
PUBMED 12477932
REFERENCE 2 (bases 1 to 4799)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (09-JAN-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library Preparation: M. Bento Soares, University of Iowa
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: University of Iowa, Dr. M. Bento Soares and Dr. Thomas L. Casavant.
Web site: <http://genome.uiowa.edu>
Contact: bento-soares@uiowa.edu; tom-casavant@uiowa.edu
Bonald, M.F., Akabogu, I., Bair, T., Bair, T., Crouch, K., Davis, A., Fishler, K., Kappel, C., Kucaba, T., Lebeck, M., Melo, A., Schaefer, K., Scheetz, J., Smith, C., Snir, E., Tack, D., Trout, K., Walters, J., Casavant, T., Soares, M.B.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: Plate: Row: Column: 0
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 27229207.
FEATURES
Location/Qualifiers
1..4799
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="CS7BL/6"
/db_xref="taxon:10090"
/clone="MGC:58057 IMAGE:6410075"
/tissue_type="Brain, enriched mouse brain 12.5dp"
/clone_lib="NIH BMAP_F00"
/lab_host="DH10B"
/note="Vector: pYX-ASC"
1..4799
/gene="Dpp8"
/db_xref="GeneID:74388"
/db_xref="MGI:1921638"
288..2966
/gene="Dpp8"
/codon_start=1
/product="dipeptidylpeptidase 8"
/protein_id="AAH43124.1"
/db_xref="GI:27695450"
/db_xref="GeneID:74388"
/db_xref="MGI:1921638"
CDS
1..4799
translation="WKIPSGRCNMAAMETEOLGVIFETAECEGNGESQDRPKLEP
FVVERYSQQLKLLADTRKYHGMMAKAPDFVKRTDPDGHSDRVVYILAMSGEN
RENTLFSEIPIKTINRAVLMLSHKPLLDLPQATLDYGMYSREERLEPERKIGTVGI
AAYDHPGSGTFLFOAGSGIYHIXDGGPHGFTQOPLRNLVETSCNINRMDPKLCPAD
PWIAFTHSDIWI SNLVTRERITTYVHNELANMEEDPRSAVTFVLQDEEFDYSG
YVWCQASRTSGSKIRILYEEENDESEVEI IHVTSFMLETRADSFYRPTGTANPK
VTFKSEIVDAAAGIIDVIDKELVQPFELFEGVEYIARAGWTPGKHWISILLDRS
OTHLQIHLI SPFLIPVEDDAMDQRILIESVPDSVTLPIIYBETTDIWINIHDFV
POTHEDELEPIFASECKTGPHLYKITSLKESKYKRSGLPASPDKCPKIKESITI
TSGMEVLRHGSNIWDEARKGVIFGTSDFEHLHYVTSVANPGEVVRILTDRGYS
HSCSLRSCDFEFTISKYQNPCHVCVSLYKLSPEDDPVHKTEFWATILLDSAGPLPDY
TPPEIFSFSEDTFTGLYGLYKPHDLQGGKPYTFLFYGGPQVQVNNRFGVKYFR
LNTLASLGVVWIDNRGSCHRLKFEKAFKXMGQIEIDDOVEGLQYLASQYDFIDL
DRYGLHCHSGYGYL SLAMALQRSIDI FRVATAGAPVTLWIFDYDTGYTERVMGHEDONEQ
GYLGSVAMQAEKFPSEPNRLLHLHGLDENVHFHTSILLSLFLVRAGKPYDLQIYPO
ERHSIRVPSGEHVELHLHYLQENLGSRLAALKVI"

ORIGIN

Alignment Scores:

Pred. No.: 0 Length: 4799
Score: 4536.50 Matches: 846
Percent Similarity: 97.8% Conservative: 18
Best Local Similarity: 95.8% Mismatches: 18
Query Match: 96.5% Indels: 1
DB: 9 Gaps: 1
US-10-825-632-1 (1-882) x BC043124 (1-4799)
Qy 1 MetAlaAlaAlaMetGluThrGluGlnLeuGlyValGluLeuPheGluThrAlaAspCys 20
Db 315 ATGGCAGCAGCAATGGAAACAGAAACAGCTGGGTGTCGAGATATTGTAACCTGCAGAGTGT 374
Qy 21 GluGlu---AsnIleGluSerGlnAspArgProIlyLeuGluProPheTyrValGluArg 39
Db 375 GAGGAGGGAATGGTGAATACAGGACCGCCAACTGGAGCCTTTTATGTCGAACCG 434
Qy 40 TyrSerTrpSerGlnLeuIlyLeuAlaAspThrArgIlyVstYrHisGlyTyrMet 59
Db 435 TATTCCTGGAGTCAGCTGAATAAGCTGCTGCTGATACCAAGAAATACCAAGCTGATATG 494
Qy 60 MetAlaIyAlaProHisAspPheMetPheValIlySArgAsnAspProAspGlyProHis 79
Db 495 ATGGCTAAGCGCCACATGACTTTATGTTTGTAAAGAGGACGAGATCCAGATGGCCTCAC 554
Qy 80 SerAspArgIleTyrTyrIleuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyr 99
Db 555 TCAGACAGGGTCTATTACCTTGCCATGTCGTGGTGAACACAGAGAAATAACACTATTTTAT 614
Qy 100 SerGluIleProIlyThrIleAsnArgAlaAlaValLeuMetLeuSerTrpIlyProLeu 119
Db 615 TCTGAATCCCTTAAACCATCAACAGCAGCAGCTTTAATGCTTTCTCGAAGCCCTT 674
Qy 120 LeuAspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeu 139
Db 675 TTGGATCTTTTTCAGCGGCACACTAGACTATGGGATGATTTCTCGAGAGGAGAGTACTC 734
Qy 140 ArgGluArgIyArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySer 159
Db 735 AGAGAAAGAAAGCGCATGGAAACCGTGGGATCGCAGCTATGATTACCAAGCCAGGAAGC 794
Qy 160 GlyThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValIlySArgGlyGlyProGln 179
Db 795 GGAACATCTCTGTTTCAAGCTGGCAGCGGCAATTTACCATATCAAGAGATGAGAGCCACAT 854
Qy 180 GlyPheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArg 199
Db 855 GGATTTACACACAGCCTTTGCGGCCCAATTTAGTGGAAACTAGTTGTCCCATATACGA 914
Qy 200 MetAspProIlyLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAsp 219
Db 915 ATGGATCCAAATATGCTCTGCTGATCCAGACTGGATAGCTTTTCATTTCACAGCAATGAT 974
Qy 220 IleTrpIleSerAsnIleValThrArgGluGluArgArgLeuThrTyrValHisAsnGlu 239
Db 975 ATTTGGATATCAACCTCTTAACAGGAGAGCGGAGATACATACGTACATACACATGAG 1034
Qy 240 LeuAlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu 259
Db 1035 CTAGCCAAACATGGAAGAGATCCAGATCAGCTGGTGTGTCACCTTTGTTCTTCAAGAA 1094
Qy 260 GluPheAspArgTyrSerGlyTyrTrpTrpCysProIlySAlaGluThrTrpProSerGly 279
Db 1095 GAATTTGACAGATACCTCTGCTACTGCTGGTGGTGTCCCAAGCGGAAAGAACTCTTAGTGGT 1154
Qy 280 GlyIlyIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHis 299
Db 1155 GGTAAATCTTTAGAAATCTCTATCAAGAAATGATGAATCTGAGGTGGAGATTATTCAT 1214
Qy 300 ValThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProIlyThrGly 319
Db 1215 GTTACGTCCCCCATGTTGGAAACAGGAGGGGCGAGATTCCTTTCTGTTATCCCAAAACAGGC 1274

QY	320	ThrAlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluLysArg	339	Db	2355	GGTTATGTGGTGTGGTGATAGACAAACAGGGGATCTCTGTACCGAGGACTTAAATTTGAA	2414
Db	1275	ACAGCAACCCAAAGGTCATCTTCAAGATGTCGAGATGTGTGTGATGCTGCAGGAGGG	1334	QY	700	GlyAlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGln	719
QY	340	IleIleAspValIleAspLysGluLeuIleGlnProPheGluIleGluPheGluGlyVal	359	Db	2415	GGCGCTTTAATATATAAATGGGTCAATAGAAATCGATGATCAAGTGAAGGACTCCAG	2474
Db	1335	ATTATAGATGTCATAGATAAAGAACTGGTTCAACCTTTCCGAGATCTCTGTGGAGAGTT	1394	QY	720	TyrLeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrPsr	739
QY	360	GluTyrIleAlaArgAlaGlyTyrThrProGluGlyLysTyrAlaTyrPsrIleLeuLeu	379	Db	2475	TACCTAGCATCTCAGTATGACTTCATTGACTCGATCGATGGGCATCCACGGCTGCTCC	2534
Db	1395	GAATATATGTCAGAGCCGATGGACTCCAGAGGGAACATGCTGGTGCATCTACTA	1454	QY	740	TyrGlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAla	759
QY	380	AspArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProVal	399	Db	2535	TATGTGGCTACCTCTCCCTGATGGCAATATACAGAGTCCGATATCTTCGGGTGGCT	2594
Db	1455	GACCGTTCCTCCAGACTCACTGCAGATAGTTCTCATCTCCCTGAGTATTATCCACAGTA	1514	QY	760	IleAlaGlyAlaProValThrLeuTyrPheTyrAspThrGlyTyrThrGluArgTyr	779
QY	400	GluAspAspValMetGluArgGlnAtgLeuIleGluSerValProAspSerValThrPro	419	Db	2595	ATTGCTGGGGCCCGACCTCCCTGTGGATCTTCTATGATACAGATACACGGAGCCCTAT	2654
Db	1515	GAAGATGATGTCATGGACAGACAGACACTTATAGATGCAGTTCCTGCATCTGTGACACCA	1574	QY	780	MetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAla	799
QY	420	LeuIleIleTyrGluGluThrThrAspIleTyrIleAsnIleHisAspIlePheHisVal	439	Db	2655	ATGGGTCACTCCCTGACCAAGATGAACAGGGCTACTTACTTAGGATCTGTGGCCATGCAAGCG	2714
Db	1575	CTGATCATCTATGAAGAAACACACAGACATCTGGATTAATATATCCAGATATTTTTCATGTT	1634	QY	800	GluLysPheProSerGluProAsnArgLeuLeuLeuLeuHisGlyPheLeuAspGluAsn	819
QY	440	PheProGlnSerHisGluGluThrGluIleGluPheIlePheAlaSerGluCysLysThrGly	459	Db	2715	GAGAAAGTTCCCTCAGAACCAACCGGTACTCTCTTTCATCGATGGGTCTTGGATGAGAAT	2774
Db	1635	TTTCTCAAACTCATGAAGATGAATTTGAGTTTATTTTTCCTCTGAAATGCATAAACAGGT	1694	QY	820	ValHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyr	839
QY	460	PheArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSer	479	Db	2775	GTTCACTTTGGACACACACAGTATATTGCTGAGTTTTTTTAGTGAGGGCTGGAAGCCATAT	2834
Db	1695	TTTCTGTCATCTGTATTAATTAACATCCATCTTTTAAGGAGAGCAATATAAACGGTCCAGT	1754	QY	840	AspLeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHis	859
QY	480	GlyGlyLeuProAlaProSerAspPheLysCysProIleGlyGluIleAlaIleThr	499	Db	2835	GACTTACAGATCTATCTCTCAGGAGAGCAGACATCAGAGTTCTCTGAGTCTGGAGAACAC	2894
Db	1755	GGTGACACTCTGCCCCAAGGATTTCAAGTGCTCTATTAAGAGAAATAACATTTACC	1814	QY	860	TyrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeu	879
QY	500	SerGlyGluTyrGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArg	519	Db	2895	TATGAACTGCACCTGCTCCACTACCTTCAGGAGAACCTTCGATCGGTCTGCTCTG	2954
Db	1815	AGTGTGTAATGGGAAGTACTTGGCCGCGCATGGATCTAAATATCTGGGTGATGAAGCCAGA	1874	QY	880	LysValIle	882
QY	520	ArgLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisIleLeuTyrValVal	539	Db	2955	AAAGTGATA	2963
Db	1875	AAGCTGGGTACTTTGAAGGCACCAAAAGACTCTCTTTGGAGCATCACCCTGACGTGACC	1934	RESULT 14			
QY	540	SerTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCys	559	BC059222			
Db	1935	AGTTATGCAAAACCTTGGAGAAAGTGTGAGGCTGACTGACCGTGGCTACTCACACTCCCTGC	1994	LOCUS			
QY	560	CysIleSerGlnHisCysAspPheIleSerLysTyrSerAsnGlnLysAsnProHis	579	DEFINITION			
Db	1995	TGCTCAGCCGCGCATTTGTGACTTCTCATAGTAAGTACAGCAACAGAGAAATCCACAC	2054	BC059222			
QY	580	CysValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGlu	599	IMAGE:6410075), complete cds.			
Db	2055	TGTGTGTCCTCTCAAACTCTCAAGTCTTGAGGATGACCCAGTTCTATAAAACAAAGGAA	2114	ACCESSION			
QY	600	PheTyrAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIle	619	VERSION			
Db	2115	TTTTGGGCCACCAATTTGGATTACGAGGCTCTCTCTGACTACACCCCTCCAGAAATT	2174	KEYWORDS			
QY	620	PheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAsp	639	SOURCE			
Db	2175	TTTTTTTGAAGTACTACTGGAATTACACTGTATGGAATGTTGTTATGAGCTCATGAC	2234	ORGANISM			
QY	640	LeuGlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyProGlnValGln	659	REFERENCE			
Db	2235	CTACACCTCTGGAAGAAATACCCACCTGTGTATTATATATGTTGGTCCCGAGGTGAG	2294	AUTHORS			
QY	660	LeuValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAsnLeu	679				
Db	2295	CTGGTGAAACAATCGGTTTAAAGGAGTCAAGTATTTCCGCTGACACCTTGGCTCCCTCG	2354				
QY	680	GlyTyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGlu	699				

Butterfield, V.S., Krzywinski, M.I., Skalska, U., Smalhus, D.E.,
Scherch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
Submitted (01-OCT-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabbs-remail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
CDNA Library Preparation: M. Bento Soares, University of Iowa
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www.shgc.stanford.edu>
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IPAK Plate: 125 Row: i Column: 19
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 27229207.
Location/Qualifiers

FEATURES

source

```
1. 4799
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="MGC:66620 IMAGE:6410075"
/tissue_type="Brain, enriched mouse brain 12.5dp"
/clone_lib="NIH_BMAP_F00"
/lab_host="DH10B"
/note="vector: pTX-ASC"
1. 4799
/gene="Dpp8"
/db_xref="GeneID:74388"
/db_xref="WGI:1921638"
288..2966
/gene="Dpp8"
/codon_start=1
/product="dipeptidylpeptidase 8"
/protein_id="AAH59222.1"
/db_xref="GI:37590654"
/db_xref="GeneID:74388"
/db_xref="WGI:1921638"
/translaton="MKTPSGRCNMAAMETEQLGVEIFETACERNGESQDRPKLEP
FYVRYSWSLKLLADTRKYHGYMAKAPDFMFKRTDPDPSHSDRVVYLAGSGEN
RENTLFYSEIPKTNRAAVLMSKPLLDLFOATLDYGMYSREBELLRKRIGTVGI
AAYDHPGSGTFLPQAGSGYHIKDCGPHGFTQOPLRNLVETSCPNRMDPKLCPAD
PDWATFTHSDNIWISNLVTRERRITYVHNELANMEEDPSAGVATFVLQEFDRYSG
YVWCPQARFTPSGKILRIILEYENDESEVEIITHVSPMLETRADSPRYKGTANFK
VTFKMSIIVDAAGGIIDVIDKELVQFELFEGVEYIARAGWTPEGKHWISLLDRS
QTHIQLVLSPELFIPEVDDMDRQRLIESVPSVTPLIIVETTDIWINTHDIHFVE
POTHEDEIRTPASECKTGRHLYKITSILKSKYKRSGGI.PAPSPKCPIKEIT
TSGWEVIEGRGNSIWDREARKLVFEFTKDSPLEHLYVTSYANPGEVRLTRGVS
HSCCLFSCHDFFTSKYSNQNCHVSLXSLSPEDDPVHKTFEFAWTLIDAGSLPDY
TPPFI8FSESTGTFLYMLYKPHDLQPKGYTFLVIFYGQVQVNNRFGKVKYFR
LNTLASLYVWVVDNRGCHRGKPEGAFKYMGQIIEIDQVGLQYLAQYDFIDL
DRVGHGWSYGVYLSLMAQMSDIERVALAGAPVTLMTFYDTGYTERYMGHPQNEQ
GYILGSVMAQAEKFPSEPNRLLLHGFLENVHFAHTSILLSFLVRACKPYDLQIYPQ
ERHSIRVPESGEHYELHLHLQENLGSRIAALKVI"
```

gene

CDS

ORIGIN

Alignment Scores:

Pred. No.:	0	4536.50	Length:	4799
Score:	97.8%	Conservative:	846	
Best Local Similarity:	95.8%	Mismatches:	18	
Query Match:	96.5%	Indels:	1	
DB:	9	Gaps:	1	

US-10-825-632-1 (1-882) x BC059222 (1-4799)

Qy	1	MetAlaAlaAlaMetGluThrGluLeuGlyValGluLeuPheGluThrAlaAspCys	20
Db	315	ATGGCAGCAGCAATGGAAACAGAACAGACGCTGGTGTGAGATATTTCAGAACTGCAGAGTGT	374
Qy	21	GluGlu---AsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArg	39
Db	375	GAGGAGGGAATGGTGAATCAGAGGACCGCCCAACTGGAGGCTTTTATGTGCGAACCG	434
Qy	40	TyrSerTspSerGlnLeuLysLeuAlaAspThrArgLysTyrHisGlyTyrMet	59
Db	435	TATTTTGGAGTACGTGAAAAAGCTGCTGCTGATACAGAAAAATACACAGGCTACATG	494
Qy	60	MetAlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHis	79
Db	495	ATGGCTAAGGCGCCACATGACTTTTATGTTGTAAGAGGACCGATCCAGATGGGCTCAC	554
Qy	80	SerAspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyr	99
Db	555	TCAGACAGGCTCTATTACCTTGGCCATGTCTGGTGAAGAACAGAAAAATACACTATTTAT	614
Qy	100	SerGluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTyrLysProLeu	119
Db	615	TCGTAAATCCCTTAAACCATCAACAGACGACGACTCTTAATGCTTCTCGAAGCCCCCT	674
Qy	120	LeuAspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeu	139
Db	675	TTGGATCTTTTCAGGCGACACTAGACTATGGGATGTTCTCGAGAGGAAGAGCTACTC	734
Qy	140	ArgGluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySer	159
Db	735	AGAGAAAGAAAGCGCATTTGGAACCGTGGGATCGCAGCTTATGATTACCAACCCAGGAAGC	794
Qy	160	GlyThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGln	179
Db	795	GGAACTTCCTGTTTCAAGCTGGGCGGCACTTACCATATCAAGAACTAGTTGTCACAGCAATGAT	854
Qy	180	GlyPheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArg	199
Db	855	GGATTTACACACAGCCTTTGCGGCCCAATTTAGTGGAACTAGTTGTCACAGCAATGAT	914
Qy	200	MetAspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAsp	219
Db	915	ATGGATCCAAATATATGCTGCTGATCCAGACTGGATAGCTTTCATTACAGCAATGAT	974
Qy	220	IleTrpIleSerAsnIleValThrArgGluArgLeuThrTyrValHisAsnGlu	239
Db	975	ATTGCGATATCAACCTGCTAACCCAGGAGCGGAGGATCATCATACGTACACAAATGAG	1034
Qy	240	LeuAlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu	259
Db	1035	CTAGCCACATCGAAGAGGATCCAGATCAGCTGGTGTGCCACCTTTGTTCTTCAAGAA	1094
Qy	260	GluPheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGly	279
Db	1095	GAATTTGACAGATCTCTGGCTACTGGTGTGTCCTCCCAAGCGGAAAGAACTCTAGTGT	1154
Qy	280	GlyLysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHis	299
Db	1155	CGTAAATTTCTAGAAATCTCTATGAAGAAATGATGAATCTGAGGTGGAGATTATTCAT	1214
Qy	300	ValThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGly	319
Db	1215	GTTACGTCCCCCATGTTGGAACCAAGGAGGCGAGATTCTTCTGTTATCCCAAAACAGGC	1274

QY 42 TrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMetAla 61
Db 3 TGGAGTCAGCTTAAAGAGCTGCTTGCAGATACAGAGAAATATCATGGCTACATGATGGCT 62
QY 62 LysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSerAsp 81
Db 63 AAGGCACACATGATTTTATGTTGTGAAGAGAAATGATCCAGATGGACCTCATTCAGAC 122
QY 82 ArgGlyLeuTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSerGlu 101
Db 123 AGAATCTATTACCTTGCATGCTCTGGTGAGACAGAGAAATATACACTGTTTATTCTGAA 182
QY 102 IleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeuAsp 121
Db 183 ATTCCCAAAATATCAATAGACAGCAGCTTAAATGCTCTCTTGGAGGCTCTCTTTGGAT 242
QY 122 LeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArgGlu 141
Db 243 CTTTTTCAGGCAACACTGGACTATGGAATGTATTCGAGAGAAAGAACTATTAAGAGAA 302
QY 142 ArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGlyThr 161
Db 303 AGAAACCATTTGGACAGTCGGAATGCTTCTTACGATTTATCCACAGAAAGTGAACA 362
QY 162 PheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGlyPhe 181
Db 363 TTTCTGTTTCAAGCGGTAGTGGAAATTTATCAGTAAAGATGGAGGCGCACAGGATTT 422
QY 182 ThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMetAsp 201
Db 423 AGCAACAACCTTTAAGGCCCAATCTAGTGGAACTAGTTGTCCCAACATAGCGATGGAT 482
QY 202 ProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIleTrp 221
Db 483 CCAAAATATGCGCTGCTGATCCAGACTGGATTCCTTTATATCATAGAACAGATTTGG 542
QY 222 IleSerAsnIleValThrArgGluGluArgGluLeuThrTyrValHisAsnGluLeuAla 241
Db 543 ATATCTCAATCGTAAACAGAGAGAAAGAGAGCTCACTTATGTGCACAAATGAGTAGCC 602
QY 242 AsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluPhe 261
Db 603 AACATGGAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAAGAAGATTT 662
QY 262 AspArgTyrSerGlyTyrTrpTyrCysProLysAlaGluThrThrProSerGlyGlyLys 281
Db 663 GATAGATATTCTGGCTATTGGTGTGTCCAAAGCTGAAACAACTCCCACTGGTGATAA 722
QY 282 IleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisValThr 301
Db 723 ATCTTAGAATTTCTATATGAAGAAATGATGATCTGAGGTGGAAATATTATCATGTACA 782
QY 302 SerProMetLeuGluThrArg-ArgAlaAspSerPheArgTyrProLysThrGlyThrAl 321
Db 783 TCCCTATGTTGGAAACAAGGCAGGCGAGATTTCATCCGTTATCCTTAAACACAGGTACAGC 842
QY 321 aAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIleI 341
Db 843 AAATCTTAAGTCACTTTTAAAGATGTCAGAAATATGATGATGCTGCTGAAGAGGATCAT 902
QY 341 eAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGluTyr 361
Db 903 AGATGTCTAGATATAGGAACATAATCAACCTTTTGGATTTCTATTTTGAAGGATTTGANA 962
QY 361 rIleAlaArgAlaGlyTyrThrProGluGlyLys- - - -TyrAlaTrpSerIleLeuLeuA 380
Db 963 TATTGCCAGCTGGATGGACTCTCTGAGGGAATAGTGTGCTTGTCTCCACTACTAG 1022
QY 380 sArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValG 400
Db 1023 ATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAAATTTATTTATCCCACTAG 1082

QY 400 luAspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProL 420
Db 1083 AAGATGATGTTATGAAGAGCAGAGACTCATTTAGTCAGTCCTGATTTCTGTACGCCAC 1142
QY 420 euIleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValP 440
Db 1143 TAAATATCTATGAAGAAACACAGACATCTGAGATAAATATCCATGACATCTTTTCATGTTT 1202
QY 440 heProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysGlyThrGlyP 460
Db 1203 TTTCCCAAGTCACGAGAGGAAATTCAGATTTATTTTGGCTCTGATGATGACAAACAGGTT 1262
QY 460 heArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerG 480
Db 1263 TCCGTCATTTATACAAAATTTACATCTATTTTAAAGGAAAGCAATAATAACGATCCAGTG 1322
QY 480 lyGlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrS 500
Db 1323 GTGGGCTGCTGCTCCNAGTGATTTCAAGTGTCTCTATCAAGAGAGAGATAGCAATTACCA 1382
QY 500 erGlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgA 520
Db 1383 GTGGTGAATGGGAAGTCTTGGCCGCGCATGGATCTAATATATCCAAGTTGATGAAGTCAGAA 1442
QY 520 rGluValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValS 540
Db 1443 GCGTGTGTATTTTGAAGGCCACCAAGACTCCCTTTTAGAGCATCACTGTAGCTAGTCA 1502
QY 540 erTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysC 560
Db 1503 GTTACGTAATCTCGAGAGGTGACNAGGCTGACTGACCGTGGCTACTACATCTTGTCT 1562
QY 560 ysIleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisC 580
Db 1563 GCATCAGTCAGCACTGTGACTTCTTTATAAGTAAGTATAGTAACAGAGAGATCCACACT 1622
QY 580 ysValSerLeuTyrLysLeuSerSerProGluLeuAspAspProThrCysLysThrLysGluP 600
Db 1623 GTGTGTCTTACAAAGCTATCAAGTCTCTGAAGATGACCCCACTTGCACAAACAAAGGAAT 1682
QY 600 heTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIleP 620
Db 1683 TTTGGGCCCACTTTTGGATTCAGCAGGTCTCTTCTCTGACTATCTCTCTCCAGAAATTT 1742
QY 620 heSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspL 640
Db 1743 TCTCTTTTGAAGTACTACTGGATTTACATTTGATGGATGCTCTCAACAGCCTCATGATC 1802
QY 640 euGlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnL 660
Db 1803 TACAGCTCGAAAGAAATATCTCTACTGTGCTGTTTCATATATATGGTGGTCTCAGGTGCA 1862
QY 660 euValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuG 680
Db 1863 TGTGGAATATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCCCTAGCTCTCTAG 1922
QY 680 lyTyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluG 700
Db 1923 GTTATGTGTTGTAGTAGACACAGGGGATCTCTGTCCAGGGCTCTAAATTTTGAAG 1982
QY 700 lyAlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnT 720
Db 1983 GCGCCTTTAAATATAAATGGGTCAATAGAAATAGACATCAGGTGGAAGGATCCCAAT 2042
QY 720 yrlLeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerT 740
Db 2043 ATCTAGCTTCTCGATATGATTTTCATTTAGTCTGTGGGATCCACGGCTGGTCTCT 2102
QY 740 yrlGlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaI 760
Db 2103 ATGGAGGATACCTCTCTCCCTGATGGCATTTATGCAGAGGTGAGATATCTTCAGGGTGTCTA 2162
QY 760 leAlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrM 780

Db	2163	TTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGATACACGGAAACGTTATA	2222
Qy	780	etGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaG	800
Db	2223	TGGGTACCCCTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCCAAGCAG	2282
Qy	800	luLysPheProSerGluProAsnArgLeuLeuLeuLeuHisGlyPheLeuAepGluAeny	820
Db	2283	AAAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGGTTTCTGGATGAGAAATG	2342
Qy	820	alHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrA	840
Db	2343	TCCATTTTGCACATACCAGTATATTACTGAGTTTTTTAGTAGGGCTGGAAAGCCATATG	2402
Qy	840	epLeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHist	860
Db	2403	ATTTACAGATCTATCTCTCAGGAGACACACAGCATNAGAGTTCTGAAATCGGGAGAACATT	2462
Qy	860	yrGluLeuHisGluLeuHisTyrTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeuL	880
Db	2463	ATGAACCTGCATCTTTTGCACTACCTTCAGAAAACCTTGGATCAGGTATTGCTGCTCTAA	2522
Qy	880	ysValIle	882
Db	2523	AAGTGATA	2530

Search completed: April 15, 2006, 03:40:39
Job time : 11208 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 15, 2006, 00:03:07 ; Search time 1249 Seconds

(without alignments)
4706.372 Million cell updates/sec

Title: US-10-825-632-1

Perfect score: 4700

Sequence: 1 MAAAMETQLGVETATDC.....HLLHYLQENLGSRIAALKVI 882

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4996997 seqs, 332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US10825632/runat_14042006_092558_4645/app_query.fasta.1
-DB=N_Geneseq -QMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss08
-USER=US10825632 @CGN_1_1_727 @runat_14042006_092558_4645 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N_Geneseq_21.*
1: Geneseqn1980s.*
2: Geneseqn1990s.*
3: Geneseqn2000s.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002as.*
7: Geneseqn2002bs.*
8: Geneseqn2003as.*
9: Geneseqn2003bs.*
10: Geneseqn2003cs.*
11: Geneseqn2003ds.*
12: Geneseqn2004as.*
13: Geneseqn2004bs.*
14: Geneseqn2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4700	100.0	2649	ABX12255	ABX12255 cDNA enco
2	4700	100.0	2649	ADV43981	Adv43981 Human psy
3	4700	100.0	2671	ABK83322	Abk83322 cDNA enco
4	4700	100.0	3106	ABK12892	Abk12892 Human pro

5	4700	100.0	3120	4	AAC85694	Aac85694 Nucleotid
6	4700	100.0	3120	6	AAD38956	Ad38956 Human dip
7	4700	100.0	3143	6	AAH99934	Aah99934 cDNA enco
8	4695	99.9	2643	6	AAH99935	Aah99935 Coding se
9	4680	99.6	4829	6	ABK83327	Abk83327 cDNA enco
10	4443.5	94.5	2797	12	ADL13374	Adl13374 Human ste
11	4397.5	93.6	2952	10	ACA92421	Aca92421 DNA enco
12	4385.5	93.3	4685	6	ABK83332	Abk83332 cDNA enco
13	4385	93.3	4676	6	ABK83331	Abk83331 cDNA enco
14	4353	92.6	2929	6	ACA92425	Aca92425 DNA enco
15	4118	87.6	2349	14	ADV43982	Adv43982 Human psy
16	4118	87.6	2842	6	ABN59774	Abn59774 Novel hum
17	4092.5	87.1	4523	6	ABK83325	Abk83325 cDNA enco
18	3970.5	84.5	2510	6	AD23843	Ad23843 Human pro
19	3901	83.0	2896	12	ADI16386	Adi16386 Human pro
20	3817	81.2	2702	13	ADT04072	Adt04072 Human pro
21	3771	80.2	2668	6	ABN59775	Abn59775 Novel hum
22	3661.5	77.9	4309	6	ABK83328	Abk83328 cDNA enco
23	3641	77.5	2251	10	ABE79035	Ade79035 Human pro
24	3597.5	76.5	2463	10	ACA92424	Aca92424 DNA enco
25	3364.5	71.6	2161	4	AAH15009	Aah15009 Human CD-
26	3087	65.7	2083	7	ADR1222	Adr1222 Human CD-
27	2870	61.1	2617	6	ABK83323	Abk83323 cDNA enco
28	2870	61.1	3716	6	ABQ75955	Abq75955 Human PMW
29	2870	61.1	4147	12	ADQ8753	Adq8753 Human tum
30	2870	61.1	4147	13	ADQ84273	Adq84273 Human tum
31	2870	61.1	4147	13	ACN40423	Acn40423 Tumour-as
32	2870	61.1	4219	6	ABK83335	Abk83335 cDNA enco
33	2870	61.1	4302	6	ABK83333	Abk83333 cDNA enco
34	2870	61.1	4398	13	ADS10267	Ads10267 Human the
35	2863	60.9	3024	6	AAD38954	Ad38954 Human dip
36	2862	60.9	2660	6	ADI16689	Adi16689 Human NOV
37	2862	60.9	2660	6	ADI16687	Adi16687 Human dip
38	2835	60.3	2495	6	AAD38957	Ad38957 Human dip
39	2833	60.3	3287	6	AAD38955	Ad38955 Alternati
40	2832	60.3	3287	6	AAD38955	Ad38955 Alternati
41	2820.5	60.0	2660	12	ADN42341	Adn42341 Human CDN
42	2820.5	60.0	4263	6	ABK83339	Abk83339 cDNA enco
43	2820.5	60.0	4359	13	ADS10269	Ads10269 Human the
44	2801	59.6	1821	6	ABV76411	Abv76411 Dipeptidy
45	2763	58.8	2751	6	AAD38311	Aad38311 Murine di

ALIGNMENTS

RESULT 1
ABX12255
ID ABX12255 standard; cDNA; 2649 BP.
XX
AC ABX12255;
XX
XX
DT 19-MAY-2003 (first entry)
XX
DE cDNA encoding human serine protease HIPHUM46.
XX

Human; ss; gene; HIPHUM46; serine protease; gene therapy; osteoarthritis;
serine protease activity modulation; dipeptidyl peptidase activity;
musculoskeletal disease; Hepatitis B virus infection; myotonic dystrophy;
Alzheimer's disease; parasupranuclear palsy; Huntington's disease;
amyotrophic lateral sclerosis; malabsorption syndrome; lung disease;
irritable bowel syndrome; type I diabetes; faecal incontinence; anaemia;
haemorrhoid; proctitis; rectal polyp; small bowel tumour; dyslexia;
colorectal tumour; ceroid lipofuscinosis; allergic encephalomyelitis;
multiple sclerosis; chromosome 15q21-q22.

XX Homo sapiens.

XX Key Location/Qualifiers
CDS

FT

FT

FT

XX

PN GB2374869-A.
 XX 30-OCT-2002.
 PD 22-JAN-2002; 2002GB-00001404.
 PF 23-JAN-2001; 2001GB-00001760.
 PR (GLAXO) GLAXO GROUP LTD.
 PA Edbrooke MR, Lewis AP;
 PI WPI: 2003-150703/15.
 DR P-PSDB; ABU07720.
 XX Identifying modulators of serine protease activity useful for treating
 PT musculoskeletal diseases, by contacting cell expressing a novel serine
 PT protease polypeptide with a compound and monitoring serine protease
 PT activity.
 XX
 PS Claim 12; Page 22-26; 38pp; English.
 XX The invention relates to a method of identifying a substance that
 CC modulates serine protease activity, comprising contacting a cell such as
 CC a neuronal cell, lung cell, intestinal cell or a cell infected with a
 CC virus, expressing a serine protease polypeptide (HirpUM 46), or its
 CC variant having dipeptidyl peptidase activity, or a serine protease
 CC isolated from the cell with a test substance and monitoring for serine
 CC protease activity. The method is useful for identifying a substance that
 CC modulates serine protease activity. A modulator of the serine protease is
 CC useful in the manufacture of a medicament for treatment or prophylaxis of
 CC a musculoskeletal disease e.g. osteoarthritis, Hepatitis B virus
 CC infection, Alzheimer's disease, paraneoplastic palsy, myotonic
 CC dystrophy, Huntington's disease or amyotrophic lateral sclerosis.
 CC Additional disease that may be treated using modulators of the serine
 CC protease include malabsorption syndromes, irritable bowel syndrome, lung
 CC disease, type I diabetes, faecal incontinence, haemorrhoids, proctitis,
 CC rectal polyps, small bowel tumours, colorectal tumours, anaemia,
 CC dyselexia, ceroid lipofuscinosis, allergic encephalomyelitis, and multiple
 CC sclerosis. The present sequence represents cDNA encoding the human serine
 CC protease HirpUM46 which is located on chromosome 15q21-q22
 XX
 SQ Sequence 2649 BP; 803 A; 514 C; 585 G; 747 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 0 Length: 2649
 Score: 4700.00 Matches: 882
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 8 Gaps: 0

US-10-825-632-1 (1-882) x ABX12255 (1-2649)

Qy 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluLeuPheGluThrAlaAspCys 20
 Db 1 ATGGCAGCAGCAATGGAAACAGAAACAGCTGGGTGTGAGATATTGGAACCTGGGACTGT 60
 Qy 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
 Db 61 GAGGAGAATATTGAATACAGAGATCGGCTAAATGGAGCCCTTTTATGTGGCCGGTAT 120
 Qy 41 SerTrpSerGlnLeuLysLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
 Db 121 TCCTGGAGTCAGCTTAAAGAGCTGTCTCCGATACCAAGAAATATCATGCGCTACATGATG 180
 Qy 61 AlaValAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
 Db 181 GCTAAGGCCACCATGATTTTCATGTTTGTGAAGAGGAATATCATGATGACCTCATTTCA 240
 Qy 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
 Db 241 GACAGAATCTATTACCTTCCCATGCTGGTGAGACAGAGAAATACACATGTTTATTCT 300

Qy 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
 Db 301 GAAATTCCTCAAACTATCAATAGACGACAGTCTTAAATGCTCTCTTGAAGCCTCTTTTG 360
 Qy 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140
 Db 361 GATCTTTTTCAGGCACACAGTGGACTATGGAAATGATTTCTCGAGAAGAGAACTATTAGA 420
 Qy 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
 Db 421 GAAAGAAAAACGATTTGGAACAGTTCGAAATTCCTTACGATTATCAACAGAGATGGA 480
 Qy 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
 Db 481 ACATTTCTGTTCGAGCGGTAGTGGAAATTTATCAAGTAAAGATCGAGGGCCACAAAGA 540
 Qy 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
 Db 541 TTTAGGCAACACCTTTAAGGCCCAATCTAGTGGAAACTAGTTGTCCCAACATACGATG 600
 Qy 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
 Db 601 GATCCAAATATTATGCCCTGCTGATCCAGACTGGATTGCTTTATACATAGCAACGATATT 660
 Qy 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu 240
 Db 661 TGGATATCTAACATCGTAACCCAGAGAAAGAGAGACTCACTTATGTGCACATAGACTA 720
 Qy 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu 260
 Db 721 GCCACATGGGAAGAGATGCCAGATCAGTGGAGTGGTACCTTTGTTCTCCCAAGAGAA 780
 Qy 261 PheAspArgTyrSerGlyTyrTrpCysProLysValGluThrThrProSerGlyGly 280
 Db 781 TTGTATAGATATTCTGGCTATTGGTGTGTCCTCAAAAGCTGAAACCACTCCAGTGGTGT 840
 Qy 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHisVal 300
 Db 841 AAAATCTTTAGAAATTTCTATATGAGAAAAATGATGATCTGAGGTGGAAATATTATCATGT 900
 Qy 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320
 Db 901 ACATCCCTTATGTTGGAACCAAGGAGGCGAGATTCTTCGTTATCTTAAACAGGTACA 960
 Qy 321 AlaAsnProLysValThrPheLysMetSerGluLeuMetIleAspAlaGluGlyArgIle 340
 Db 961 GCAATCTCTAAAGTCACTTTTAAAGATGTGAGAAATAATGATTGATGCTGAAGGAAGATC 1020
 Qy 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
 Db 1021 ATAGATGTCTATAGATAAGGAACCTAATTCAACCTTTTGAGATTCTATTGAGGGAGTTGAA 1080
 Qy 361 TyrIleAlaArgAlaGlyTyrThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
 Db 1081 TATATTGCCAGAGCTGGATGGACTCTCGAGGAAAAATATGCTTGGTCCATCTACTAGAT 1140
 Qy 381 ArgSerGlnThrArgGluGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
 Db 1141 CGCTCCAGACTCGCTCGAGATAGTGTGATCTCACCTGAATATTATTATCCCAAGTAGAA 1200
 Qy 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
 Db 1201 GATGATGTTATGGAAGGCGAGACTCATTTGATCAGTGCCTGATTCTGTGAGGCCACTA 1260
 Qy 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
 Db 1261 ATTATCTATGAAGAAAAACACACATCTGGATTAATATCCATGACATCTTTCATGTTTT 1320
 Qy 441 ProGlnSerHisGluGluGluLeuGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
 Db 1321 CCCCAGAGTCACAGAGGAAAAATTGAGTTTATTTTGGCTCTGAAATGCAAAACAGGTTTC 1380

Qy	1	MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys	20	Db	1081	TATATTCCAGAGCTGGATGGACTCTCTGAGGGAAATATATGCTTGGTCCACTACTAGAT	1140
Db	1	ATGGCAGCAGCAATGGAAACAGAACAGCTGGGTGTGTAGATATTTGAAACATGGCGGACTGT	60	Qy	391	ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu	400
Qy	21	GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr	40	Db	1141	CGCTCCAGAGCTGGCTTACAGATAGTGTGTGATCTACCTGAATATATTTATCCAGTAGAA	1200
Db	61	GAGGAGAAATATTGAATCACAGGATCGGCTTAAATTGGAGCCCTTTTATATGTTGAGCGGTAT	120	Qy	401	AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu	420
Qy	41	SerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet	60	Db	1201	GATGATGTTATGGAAAGGAGAGACTCATTTGAGTCAGTCCCTGATCTCTGTGACGCACATA	1260
Db	121	TCCTGGATCAGCTTAAAGAGCTGCTTGGCGATACCAGAAAAATATCATGGCTACATGATG	180	Qy	421	IleIleTyrGluGluThrThrAspIleTyrIleAsnIleHisAspIlePheHisValPhe	440
Qy	61	AlaTyrAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer	80	Db	1261	ATTATCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTTTCTATGTTTTT	1320
Db	181	GCTAAGGCCACCATGATTTTCATGTTTGTGAAGAGAAATGATCCAGATGGACCTCATTTCA	240	Qy	441	ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe	460
Qy	81	AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer	100	Db	1321	CCCCAAAGTCACGAAGAGGAAATTGAGTTTTATTTTTGCTCTGAAATGCAAAACAGGTTTC	1380
Db	241	GACAGAACTCTATTACCTTGGCATGTCTGGTGAGAACAGAGAAAATACACTGTTTATTTCT	300	Qy	461	ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly	480
Qy	101	GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu	120	Db	1381	CGTCATTTATACAAATTTACATCTATTTTAAAGAAAGCAAAATATAAAGCATCCAGTGGT	1440
Db	301	GAATATCCCAAAACTATCAATAGAGCAGCAGCTCTTAATGCTCTCTTGGAGCCCTCTTTTG	360	Qy	481	GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer	500
Qy	121	AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg	140	Db	1441	GGGCTGCTGCTCCCAAGTGATTTCAAGTGTCTTATCAAAGAGGAGATAGCAATTTCCAGT	1500
Db	361	GATCTTTTTCAGGCAACACTGGACTATGGAAATGTATTTCTCGAAGAGAAAGACTATTAGA	420	Qy	501	GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg	520
Qy	141	GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly	160	Db	1501	GGTGAAATGGGAAGTCTCTGGCCGCGATGGATCTAATATCAAGTTTGATGAAGTCAGAAAG	1560
Db	421	GAAGAAAAACGCATTTGGAAACAGTCGGAAATTTCTTTCATCGATTTATCACCAAGAAAGTGA	480	Qy	521	LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSer	540
Qy	161	ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly	180	Db	1561	CTGGTATATTTTGAAGGACCAAGAGCTCCCTTTAGAGCATCACCTGTACCTAGTTCAGT	1620
Db	481	ACATTTCTGTTTCAAGCCGGTAGTGGAAATTTATACGTTAAAGATGGAGGGCCACAAAGA	540	Qy	541	TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys	560
Qy	181	PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet	200	Db	1621	TACGTAATCTCTGGAGAGTGACAAAGCTGACTGACCGTGGCTACTCACATTTCTTGCTGC	1680
Db	541	TTTACGCAACACCTTTTAAGCCCAATCTAGTGGAAATAGTTGTCCCAACATACGGATG	600	Qy	561	IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys	580
Qy	201	AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle	220	Db	1681	ATCAGTCAGCACTGTGACTCTTTATAGTAAGTATAGTAACCAAGAAAGAAATCCCACTGT	1740
Db	601	GATCCAAATAATATGCTCCCTGCTGATCCAGACTGGATGCTTTTATACATAGCAACGATAT	660	Qy	581	ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe	600
Qy	221	TrpIleSerAsnIleValThrArgGluGluArgArgLeuThrTyrValHisAsnGluLeu	240	Db	1741	GTGTCCCTTTTACAAAGTATCAAGTCTCTGAAGATGACCCAACTTGCAAAACAAAGGAATTT	1800
Db	661	TGATATCTTACATCTGTAAACAGAGAGAAAGAGAGACTCACTTATGTGCACATAGACTA	720	Qy	601	TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe	620
Qy	241	AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu	260	Db	1801	TGGGCCACCATTTTGGATTCAGCAGGTCCTCTTCTGACTATATCTCTCAGAAAATTTTC	1860
Db	721	GCCAACTGGAAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAGAGAGAA	780	Qy	621	SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu	640
Qy	261	PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly	280	Db	1861	TCCTTTGAAAGTACTACTTGGATTTTACATTTGTATGGGATGCTCTCAAGCCTCATGATCTA	1920
Db	781	TTTGTAGATATATCTGGCTATTTGGTGGTGTCCAAAGCTGAACACACTCCAGTGGTGT	840	Qy	641	GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu	660
Qy	281	LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal	300	Db	1921	CAGCTCGAAAGAAATATCTTACTGTGCTGTTTATATATATATGTTGCTCTCAGGTGCGAGTTG	1980
Db	841	AAAAATCTTAGAATTTCTATGAGAAAAATGATGAATCTGAGGTGGAATATTTATCAATGT	900	Qy	661	ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly	680
Qy	301	ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr	320	Db	1981	GTGAATATTCGGTTTAAAGGAGTCAAGTATTTTCCGTTTGAATACCCCTAGGCTCTCTAGGT	2040
Db	901	ACATCCCTCATGTGTGAAACAGAGGGGAGAGATTCATTCGGTATATCTCTAAACAGAGTACA	960	Qy	681	TyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly	700
Qy	321	AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluArgIle	340	Db	2041	TATGTGTTTGTAGTGTATAGACAAACAGGGGATCCTGTCCACGAGGGCTTAAATTTGAAGGC	2100
Db	961	GCAATCTTAAAGTCACCTTTTAAAGATGTGAGAAATAATGATGTGATGCTGAAGGAAGGATC	1020	Qy	701	AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr	720
Qy	341	IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu	360	Db	2101	GCCTTTAAATATATAAATGSGTCAATAGAAATTCAGCATCAGGTGGAAGGACTCCAATAT	2160
Db	1021	ATAGATGTCATAGATAAGGAACATAATCAACCTTTTGAGATTTCTATTTTGAAGGAGTTGAA	1080	Qy	721	LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyr	740
Qy	361	TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp	380				

Db 2161 CTAGCTTCGATGATGATTCATTGACTTTAGATCGTGTGGGCATCCACGGCTGTCCTAT 2220

Qy 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760

Db 2221 GGAGGATACCTCTCCCTGATGGCATTAATGACAGGTGAGATATCTTCAGGGTTCCTATT 2280

Qy 761 AlaGlyAlaProValThrLeuThrPheTyrAspThrGlyTyrThrGluArgTyrMet 780

Db 2281 GCTGGGGCCCCAGTCACCTCTGTGGATCTTCTATGATACAGGATACACAGGAACGTTATATG 2340

Qy 781 GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800

Db 2341 GGTCAACCTGACCAAGTAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGAGAA 2400

Qy 801 LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820

Db 2401 AAGTTCCTCCCTTGAAACCAATCGTTTACTGCTCTTACATGGTTCTCTGATGAGATGTC 2460

Qy 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840

Db 2461 CATTTTGCATACACAGATATATTACTAGTTTTTGTAGGGCTGGAAAGCCATATGAT 2520

Qy 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860

Db 2521 TTACAGATCTATCTCTCAGGAGACACAGCATAGAGTTCCTGATCGGGAGACATTAT 2580

Qy 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLys 880

Db 2581 GAACTGCACTCTTTTGCACCTACCTTCAAGAAACCTTGGATCAGTATTGCTGCTTAAGA 2640

Qy 881 ValIle 882

Db 2641 GTGATA 2646

RESULT 3

ABK83322

ID ABK83322 standard; cDNA; 2671 BP.

XX AC ABK83322;

XX

DT 12-AUG-2002 (first entry)

XX cDNA encoding human DPPIV related serine protease DPPP-1.

DE

XX Human; serine protease; dipeptidyl peptidase IV-related protein; DPPP; DPPIV; infection; human immunodeficiency virus; HIV-1; HIV-2; pain; diabetes; infertility; obesity; anorexia; Parkinson's disease; stroke; heart failure; hypertension; urinary retention; osteoporosis; cancer; ulcer; allergy; cancer; psychotic disorder; neurological disorder; dyskinnesia; reproductive disorder; inflammatory disorder; metabolic disorder; gene; ss.

XX

OS Homo sapiens.

XX

PN WO200231134-A2.

XX

PD 18-APR-2002.

XX

PF 12-OCT-2001; 2001WO-US031874.

XX

PR 12-OCT-2000; 2000US-0240117P.

XX (FERR) FERRING BV.

XX

PA Qi S, Akinsanya KO, Riviere PJ, Junien J;

XX

PI WPI; 2002-444178/47.

XX

DR P-PSDB; ABG61591.

XX

XX New dipeptidyl peptidase IV-related proteins and nucleic acids encoding the proteins, useful for treating e.g. fungal, bacterial, protozoan and viral infections, cancers, allergies, neurological disorders, or pain.

XX

PS Claim 1; Page 53-54; 113pp; English.

XX The present invention relates to the isolation of novel human serine proteases referred to as dipeptidyl peptidase IV (DPPIV)-related proteins (DPPP). The dipeptidyl peptidase IV-related proteins (DPPP) and nucleic acids encoding them are useful for treating infections such as fungal, bacterial, protozoan and viral infections, particularly infections caused by human immunodeficiency virus (HIV-1 or HIV-2), pain, diabetes, disease, acute heart failure, hypotension, hypertension, bulimia, Parkinson's retention, osteoporosis, angina pectoris, stroke, ulcers, asthma, allergies, cancers, migraine, vomiting, psychotic and neurological disorders (e.g. anxiety, dementia, or schizophrenia), and dyskinesias. These may also be used in discovering therapeutic agents for the treatment of reproductive, inflammatory and metabolic disorders. ABK83322 -ABK83343 encode human DPPP proteins

SQ Sequence 2671 BP; 805 A; 524 C; 594 G; 748 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 0 Length: 2671
Score: 4700.00 Matches: 882
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0

US-10-825-632-1 (1-882) x ABK83322 (1-2671)

Qy 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20

Db 8 ATGGCAGCAGCAATGGAACACAGACAGCTGGGTGTGAGATATTGAAACTGGCGACTGT 67

Qy 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40

Db 68 GAGGAGATATTGAATCACAGATCGGCTAAATGGAGCCTTTTATGTGAGCGGTAT 127

Qy 41 SerTyrSerGlnLeuLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60

Db 128 TCCTGGAGTCAGCTTAAAAAGCTGCTGCCGATACCAAGAAATATCATGCGTACATGATG 187

Qy 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80

Db 188 GCTAAGCCACCATGATTTCATGTTGTGAAGAGGAATGATCCAGATGGACCTCATTC 247

Qy 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100

Db 248 GACAGAAATCTATTACCTTGCCATGCTGGTGAAGAACAGAGAAATACACTGTTTATCT 307

Qy 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTyrLysProLeuLeu 120

Db 308 GAATTTCCCAAACTATCAATAGACAGCAGCTTATGCTCTCTTGGAGCCTCTTTTG 367

Qy 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuArg 140

Db 368 GATCTTTTTCAGCAACACTGGACTATGGAATGATTTCTCAGAGAAGAACTATTAA 427

Qy 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160

Db 428 GAAAGAAAACGGCATTTGGAAACAGTCGGAATTTGCTTACGATTTATCCCAAGAGTGA 487

Qy 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180

Db 488 ACATTTCTTTTCAGCCCGGTAGTGAATTTATCATGTAAGATGGAGGCCACACAGGA 547

Qy 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200

Db 548 TTTACGCAACACACTTTAAGGCCAATCTAGTGGAACTAGTGTGCCCAACATACATG 607

Qy 201 AspProLysLeuCysProAlaAspProAspThrIleAlaPheIleHisSerAsnAspIle 220

Db 608 GATCAAAATATGCTGCTGATCCAGACTGGATTGCTTTTATACATAGCAAGATATT 667

QY	221	TrpIleSerAenIleValThrArgGluGluArgLeuThrTyxValHisAsnGluLeu	240	Db	1748	GTGTCCCTTTCAAGACTATCAAGTCTCGAAGATGACCCCACTTGCRAAACAAGGAATTT	1807
Db	668	TGGATATCTAAACATCGTAACAGAGAGAAGAAGAGACTCACATTATGTGACATANGCTA	727	QY	601	TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePhe	620
QY	241	AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu	260	Db	1808	TGGGCCACCATTTTGGATTTCAGCAGGTCTCTCTCTGACTATACTCTCTCAGAAATTTTC	1867
Db	728	GCCAAACATGGGAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAAGAAGA	787	QY	621	SerPheGluSerThrThrGlyPheThrLeuTyxGlyMetLeuTyxIlePheProHisAspLeu	640
QY	261	PheAspArgTyrSerGlyTyrTTPtpCyseProLysAlaGluThrThrProSerGlyGly	280	Db	1868	TCCTTTTGAAGATCTACTGGATTTACATTTGTATGGGATGCTCTACAGGCTCTCAATGATCA	1927
Db	788	TTTGTAGATATATCTCGGCTATTTGGTGTGTCCAAAAGCTGAAACAACCTCCCAAGTGGTGT	847	QY	641	GlnProGlyLysLysThrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu	660
QY	281	LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal	300	Db	1928	CAGCCTGGAAAAGAAATATCTCTACTGTCTGTTCATATATGCTGGTCTCTCAGGTGCAAGTTC	1987
Db	848	AAAAATCTTAGAAATCTATATCAAGAAAAATGATGAATCTGAGGTGGAATATATTCATGTT	907	QY	661	ValAsnAsnArgPheLysGlyValLysTyxPheArgLeuAsnThrLeuAlaSerLeuGly	680
QY	301	ThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThr	320	Db	1988	GTGAATAATTCGATTTAAAGGAGTCAAGTATTTCCGCTTGAATACCCCTAGCCTCTCTAGGT	2047
Db	908	ACATCCCCCTATGTTGGAACAAGGAGGCGAGATTCATTCCGTTATCTTAAAAACAGGTACA	967	QY	681	TyrValValValValIleAspAsnArgGlySerCyseHisArgGlyLeuLysPheGluGly	700
QY	321	AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle	340	Db	2048	TATGTGGTTGTAGTGATAGACACACAGGGGATCTCTGTCCCGAGGGCTTAAATTTGAAGGC	2107
Db	968	GCAAAATCTTAAGTCACTTTTAAGATGTCAGAAATATGATTCATGCTGAAGGAAGGATC	1027	QY	701	AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr	720
QY	341	IleAspValIleAspLysGluLeuIleGlnProPheGluIleIleuPheGluGlyValGlu	360	Db	2108	GCTTTAATAATAAATGGGTCAATAGAAATTGACGATCAGGTGGAAGGACTCCCAATAT	2167
Db	1028	ATAGATGTCATAGATAGAGAACTAAATTCAACTTTTGAGATCTATTGGAAGAGTTGAA	1087	QY	721	LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyr	740
QY	361	TyrIleAlaArgAlaGlyTyrThrProGluGlyLysTyrAlaTyrSerIleLeuLeuAsp	380	Db	2168	CTAGCTTCTCGATATGATTTTCATTTGACTTAGATCGTGTGGGCATCCACGGCTGGTCTAT	2227
Db	1088	TATATTTGCCAGAGCTGGATGACTCTCGAGGGAATAATGCTTGGTCCATCTACTAGAT	1147	QY	741	GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle	760
QY	381	ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu	400	Db	2228	GGAGGATACCTCTCCCTGATGGCATTAATGCAGAGGTCAATATCTCAGGGTTGCTATT	2287
Db	1148	CGCTCCCAAGACTCGCTCGCATAGATGTTGATCTCACTCCCTGAAATATTTATCCCAAGTAA	1207	QY	761	AlaGlyAlaProValThrLeuTyrIlePheTyrAspThrGlyTyrThrGluArgTyrMet	780
QY	401	AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu	420	Db	2288	GCTGGGGCCCCAGTCACCTCTGTGGATCTTCTATGATACAGGATACACGGAACTTATATG	2347
Db	1208	CATGATGTTATGGAAAGCAGAGACTCAATTTGAGTCAGTCGCTGATTCGTGACGCCACTA	1267	QY	781	GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu	800
QY	421	IleIleTyrGluGluThrThrAspIleTyrIleAsnIleHisAspIlePheHisValPhe	440	Db	2348	GGTCAACCTCGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGCCATGCAAGCAGAA	2407
Db	1268	ATTATCTATGAAGAACCAACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTTT	1327	QY	801	LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal	820
QY	441	ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe	460	Db	2408	AAAGTTCCCTCTGAAACCAAAATCGTTTACTGTCTTACATGGTTTCTCTGGATGAGAATGTC	2467
Db	1328	CCCCAAAGTCACGAAGAGGAATTTGAGTTTATTTTGGCTCTGAATGCAAAACAGGTTTC	1387	QY	821	HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp	840
QY	461	ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly	480	Db	2468	CATTTTGACATACCCAGTATATTTACTGAGTTTTTTAGTGAGGCTGGAAGCCATATGAT	2527
Db	1388	CGTCATTTATACAAATTTACATCTATTTTAAGGAAGCAATATATAACGATCCAGTGT	1447	QY	841	LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr	860
QY	481	GlyLeuProAlaProSerAspPheLysCyseProIleLysGluLeuIleAlaIleThrSer	500	Db	2528	TTACAGATCTATCTCCAGAGACACACAGCATTAAGAGTTCCTGAAATCGGGAGAACATTA	2587
Db	1448	GGGCTGCCTCTCCANGTGATTTCAAGTGTCTCTATCAAGAGAGATAGCAATTTACCA	1507	QY	861	GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLys	880
QY	501	GlyGluTyrGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArg	520	Db	2588	GAACTGCATCTTTTGCACTACTCTTCAAGAAAAACCTTGATTCACGTATTTGCTGCTTAAA	2647
Db	1508	GGTGAATGGGAAGTTCTTTGGCGGCATGGATCTAAATATCCAGTTGATGAAGTCAGAAG	1567	QY	881	ValIle 882	
QY	521	LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer	540	Db	2648	GTGATA 2653	
Db	1568	CTGGTATATTTTGAAGGCACCAAGACTCCCTTTTGAAGCATCACCTGTACGTAGTCAGT	1627	QY			
QY	541	TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCyse	560	Db			
Db	1628	TACGTAATCTCTGGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCACATTTCTTGCTGC	1687	Db			
QY	561	IleSerGlnHisCyseAspPheIleSerLysTyrSerAsnGlnLysAsnProHisCyse	580	Db			
Db	1688	ATCAGTCAGCACTGTGACTCTCTTTATAAGTAAAGTATAGTAACCAAGAAAGTAATCCACACTGT	1747	Db			
QY	581	ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe	600	Db			

RESULT 4

ABK12892

ID ABK12892 standard; cDNA; 3106 BP.

XX

AC ABK12892;

XX

XX 09-APR-2002 (first entry)

DT

XX Human protease PRTS-9 cDNA sequence.

DE

XX

Human; protease; PRTS; gastrointestinal; Crohn's disease; cancer; cardiovascular; atherosclerosis; autoimmune disorder; dermatitis; inflammatory disorder; acquired immunodeficiency syndrome; AIDS; cell proliferative disorder; developmental disorder; epilepsy; Duchenne muscular dystrophy; epithelial disorder; neurological disorder; reproductive disorder; endometriosis; ss.

Homo sapiens.

Key Location/Qualifiers
CDS 203..2851

/*tag= a

/product= "Human protease PRTS-9"

W0200198468-A2.

27-DEC-2001.

13-JUN-2001; 2001WO-US019178.

16-JUN-2000; 2000US-0212336P.

22-JUN-2000; 2000US-0213955P.

29-JUN-2000; 2000US-021396P.

07-JUL-2000; 2000US-0216821P.

14-JUL-2000; 2000US-0218946P.

(INCY-) INCYTE GENOMICS INC.

Yue H, Elliott VS, Gandhi AR, Lal P, Au-Young J, Tribouley CM; Deleage AM, Baughn MR, Nguyen DB, Lee EA, Hafalia A, Khan FA; Wallia NK, Yao MG, Lu DAM, Patterson C, Tang YT, Walsh RT; Azimzai Y, Lu Y, Ramkumar J, Xu Y, Reddy R, Das D, Kearney L; Kallick DA;

WPI; 2002-090437/12.

P-PSDB; AAU74749.

Twenty one human proteases (referred to as PRTS-1 to PRTS-21), useful in the diagnosis, treatment and prevention of gastrointestinal (e.g. gastritis), cardiovascular (e.g. atherosclerosis) and cell proliferative (e.g. cancer) disorders.

Claim 5; Page 166-167; 177pp; English.

The present invention relates to twenty one new human proteases, referred to as PRTS-1 to PRTS-21. The PRTS polynucleotides and polypeptides of the invention are useful in the diagnosis, treatment and prevention of gastrointestinal e.g. gastritis, esophageal carcinoma and Crohn's disease, cardiovascular e.g. atherosclerosis, hypertension and myocardial infarction, autoimmune/inflammatory e.g. acquired immunodeficiency syndrome (AIDS), allergies and osteoarthritis, cell proliferative e.g. cancer, developmental e.g. Duchenne and Becker muscular dystrophy, epithelial e.g. dermatitis, neurological e.g. epilepsy and Alzheimer's disease and reproductive e.g. infertility and endometriosis disorders. Numerous other examples of each disorder are given in the specification. The present nucleic acid sequence encodes the human protease PRTS-9 protein of the invention

SQ Sequence 3106 BP; 928 A; 633 C; 704 G; 841 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	0	Length:	3106
Score:	4700.00	Matches:	882
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	6	Gaps:	0

US-10-825-632-1 (1-882) x ABK12892 (1-3106)

Qy 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluilePheGluThrAlaAspCys 20
Db 203 ATGGCAGCAGCAATGGAAACAGACAGCTGGGTGTGGATATTGGAACTGGCGACTGT 262

Qy	21	GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr	40
Db	263	GAGGAGAAATATTGAATCAAGGATCGGCTAAATTTGGAGCCTTTTATGTTGAGCGGTAT	322
Qy	41	SerTyrSerGlnLeuLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet	60
Db	323	TCTGGAGTCAGCTTAAAGAGCTGCTGCCGATACCAAGAAATATCATGCTACATGATG	382
Qy	61	AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer	80
Db	383	GCTAAGGCACACATGATTTTCTGTTGTAAGAGGAATGATCCAGATGACCTCATTTCA	442
Qy	81	AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer	100
Db	443	GACAGAAATCTATTACCTTGCCTATGTCGTGAGAACAGAGAAATATCATGCTACATGATG	502
Qy	101	GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTyrLysProLeuLeu	120
Db	503	GAATTTCCCAAAACATATCAATAGACAGCAGCTTAAATGCTCTCTTGGAGAGCTCTTTTG	562
Qy	121	AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg	140
Db	563	GATCTTTTTCAGGCAACACTGGACTATGGAATGATTTCTCGAGAAGAAGAACTATTAAAG	622
Qy	141	GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly	160
Db	623	GAAGAAACACCAATTTGGAACAGTCGGAATTCCTTACGATATATCATCCCAAGGAAGTGA	682
Qy	161	ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly	180
Db	683	ACATTTCTCTTTTCAAGCCGCTAGTGGAAATTTATACGCTAAAGAGATGAGGAGGACCA	742
Qy	181	PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet	200
Db	743	TTTACGCAACAACTTTTAAAGGCCAATCTAGTGAACACTAGTTGCTCCACATACGATG	802
Qy	201	AspProLysLeuCysProAlaAspProAspTyrIleAlaPheIleHisSerAsnAspIle	220
Db	803	GATCCAAATTAATGCGCTGCTGATCCAGACTGGATGCTTTTATATACATAGCAACATAT	862
Qy	221	TyrIleSerAsnIleValThrArgGluArgLeuThrTyrValHisAsnGluLeu	240
Db	863	TGGATATCTAAACATCGTAAACAGAGAGAGAGACTCACTTATGTCACAAATGAGCTA	922
Qy	241	AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu	260
Db	923	GCCAAACATGGAAGAGATGCCAGATCAGCTGAGTCGCTACTTTGTTCTCCAAAGAGAA	982
Qy	261	PheAspArgTyrSerGlyTyrTyrTyrCysProLysAlaGluThrThrProSerGlyGly	280
Db	983	TTTGATAGATATTCTGGCTATTGGTGGTGTCCAAAGCTGAACAACTCCCCAGTGGTGT	1042
Qy	281	LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHisVal	300
Db	1043	AAAAATCTTAGAATCTATATGAGAAATGATGAATCTGAGGTGGAAATTTATTCTATGTT	1102
Qy	301	ThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThr	320
Db	1103	ACATCCCTTATGTTGAAACAAAGGAGGCGAGATTCATTCCTGTTATCTTAAACAGGTACA	1162
Qy	321	AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle	340
Db	1163	GCAAAATCTTAAGTCACTTTTAAGATGTGAGAAATATGATGCTGCTGAAGGAAGATC	1222
Qy	341	IleAspValIleAspLysGluLeuIleGlnPropheGluIleLeuPheGluGlyValGlu	360
Db	1223	ATAGATGTCATAGATAAGGAACCTAATTTCAACCTTTTGAGATCTTATTGAAAGGAGTTGA	1282
Qy	361	TyrIleAlaArgAlaGlyTyrThrProGluGlyLysTyrAlaTyrSerIleLeuLeuAsp	380
Db	1283	TATATTGCCAGAGCTGGATGGACTCTCTGGGGAATAATATGCTTGGTCCATCTACTAGAT	1342

QY	381	ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu	400
Db	1343	CGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATATATATCCAGTAGAA	1402
QY	401	AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu	420
Db	1403	GATGATGTATGGAAGGACAGACTCATTTGAGTCAGTCGCTGATTTCTGTGACGCCACTA	1462
QY	421	IleIleIleTyrGluGluThrThrAspIleTrrpIleAsnIleIleHisAspIlePheHisValPhe	440
Db	1463	ATTATCTATGAAGAACAAACAGACATCTGGATAAATATCATGACATCTTTTCATGTTTTT	1522
QY	441	ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysIleThrGlyPhe	460
Db	1523	CCCCAAGCTCAGGAAGGAAATTTGAGTTTATTTTGGCTCTGAAATGCAAAACAGGTTTC	1582
QY	461	ArgHisLeuTyrIleValIleThrSerIleLeuIleGluSerIleTyrIleArgSerSerGly	480
Db	1583	CGTCATTTATCAAAATATACATCTATTTTAAAGGAAGCAATATATAACGATCCAGTGGT	1642
QY	481	GlyLeuProAlaProSerAspPheIleCysProIleIleValGluIleAlaIleThrSer	500
Db	1643	GGGCTGCCTCGCTCCAGGATTTCAAGTGCTCTATCAAGAGGAGATAGCAATTTACAGT	1702
QY	501	GlyGluTrrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg	520
Db	1703	GGTGAATGGGAAGTCTTTGGCCGCGCATGGATCTAAATATCCAGTTGATGAATCAGAAGG	1762
QY	521	LeuValTyrPheGluGlyThrIleAspSerProLeuGluHisIleValValSer	540
Db	1763	CTGGTATATTTTGAAGGCACCAAGACTCCCTTTTAAAGCATCACCCTGTACGTAGTCAGT	1822
QY	541	TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys	560
Db	1823	TACGTAATTCCTGGAGAGGTGACAAAGCTGACTGACCGTGGCTACTCACATTTCTTGCTGC	1882
QY	561	IleSerGlnHisCysAspPheIleSerIleTyrSerAsnGlnIleAsnProHisCys	580
Db	1883	ATCAGTCAGCAGCTGTGCTCTTTATAAGTAAGTATAGTAACCAAGAAAGATCCACACTGT	1942
QY	581	ValSerLeuTyrIleLeuSerSerProGluAspAspProThrCysIleThrIleGluPhe	600
Db	1943	GTGTCCTTTTCAAGCTATCAAGTCTTGAAGATGACCCCACTTTGCCAAACAAAGGAATTT	2002
QY	601	TrrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe	620
Db	2003	TGGGCCACCATTTTGGATTCAGCAGGTCTCTTCTGACTATATCTCTCCAGAAATTTTC	2062
QY	621	SerPheGluSerThrThrGlyPheThrIleuTyrGlyMetLeuTyrIlePheHisAspLeu	640
Db	2063	TCCTTTTGAAGTACTACTGGATTTTACATTTGATGGGATGCTCTACAAAGCCTCATGATCTA	2122
QY	641	GlnProGlyIleGlyTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu	660
Db	2123	CAGCCTGGAAAGAAATATCCCTACTGTGCTGTTCATATATGTTGGTCTCCAGGTGCAGTTG	2182
QY	661	ValAsnAsnArgPheIleGlyValIleTyrPheArgLeuAsnThrLeuAlaSerLeuGly	680
Db	2183	GTGAATTAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCCCTAGCCTCTCTAGT	2242
QY	681	TyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuIleHisPheGluGly	700
Db	2243	TATGTGGTTGTAGTATAGACAAACAGGGGATCCTGTCCCGAGGGCTTAAATTTGAAGGC	2302
QY	701	AlaPheIleTyrIleSerGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr	720
Db	2303	GCCTTTTAAATATAAATGGGTCAAAATAGAAATTCACGATCAGGTGGAAGGACTCCCAATAT	2362
QY	721	LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrrpSerTyr	740
Db	2363	CTAGCTTCTCGATATGATTTTCATTTGATGATCGTGGGACATCCACGGCTGGTCTCAT	2422
QY	741	GlyGlyTyrIleLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle	760
Db	2423	GGAGGATACCTCTCCCTGATGGCATTAAATGCAGAGGTTCAGATATCTTCAGGGTTGCTATT	2482
QY	761	AlaGlyAlaProValThrLeuTrrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet	780
Db	2483	GCTGGGGCCCAAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGACGTTATATG	2542
QY	781	GlyHisProAspGlnAsnGluGlnGlyTyrTyrIleuGlySerValAlaMetGlnAlaGlu	800
Db	2543	GGTCACCTTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCATGCAAGCAGAA	2602
QY	801	LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal	820
Db	2603	AAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATCGTTTCTCGGATGAAATGTC	2662
QY	821	HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp	840
Db	2663	CATTTTGCACATACCAGTATATTACTGAGTTTTTTTAGTGGGCTGGAAAGCCATATGAT	2722
QY	841	LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr	860
Db	2723	TTACAGATCTATCTCTCAGGAGACACAGCATTAAGAGTTCTCTGAATCGGAGAACATTAT	2782
QY	861	GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLys	880
Db	2783	GAACTGCATCTTTTGCACCTACCTTCAAGAAAACCTTGGATCAGTATTGCTCTTAAAA	2842
QY	881	ValIle	882
Db	2843	GTGATA	2848
RESULT 5			
ID	AAC85694	standard; cDNA; 3120 BP.	
XX	AAC85694;		
DT	29-JUN-2001	(first entry)	
DE	Nucleotide sequence of human DPP8.		
KW	Human; dipeptidyl aminopeptidase; DPP8; prollyl oligopeptidase;		
KW	dipeptidyl peptidase; DPPIV; T cell; cleavage; diarrhoea;		
KW	growth hormone deficiency; glucose level; mucosal regeneration;		
KW	non-insulin dependent diabetes mellitus; glucose intolerance;		
XX	immunosuppression; ss.		
OS	Homo sapiens.		
XX	Location/Qualifiers		
FT	214..2862		
FT	/*tag= a		
FT	/product= "Human DPP8"		
XX	W0200119866-A1.		
XX	22-MAR-2001.		
XX	11-SEP-2000; 2000WO-AU001085.		
XX	10-SEP-1999; 99AU-00002762.		
XX	18-FEB-2000; 2000AU-00005709.		
XX	(UNSY) UNIV SYDNEY.		
XX	Abbott CA, Gorell MD;		
XX	WPI; 2001-281520/29.		
XX	P-PSDB; AAB47187.		
XX	New human dipeptidyl aminopeptidase (DPP8) useful for cleaving		
XX	substrates, identifying inhibitors of DPP8 catalytic activity which have		
XX	therapeutic uses, and for detecting activated T cells.		

XX

Claim 16; Fig 2; 78pp; English.

XX This sequence encodes human dipeptidyl aminopeptidase (DPP8). DPP8 has
 CC substrate specificity for H-Ala-Pro-pNA, H-Gly-Pro-pNA and H-Arg-Pro-pNA.
 CC Therefore, it is a prolyl oligopeptidase and a dipeptidyl peptidase.
 CC because it is capable of hydrolyzing the peptide bond C-terminal to Pro
 CC in each of these compounds. DPP8 is homologous with human DPPIV. DPP8 is
 CC useful for cleaving a substrate, and for detecting an activated T cell
 CC which involves measuring the level of DPP8 gene expression in a T cell.
 CC The level of DPP8 expression is detected by detecting the amount of DPP8
 CC RNA in the cell. It is also useful for identifying a molecule capable of
 CC inhibiting the cleavage of the substrate by DPP8. Molecules identified as
 CC inhibiting DPP8 catalytic activity may be useful for treating diarrhoea,
 CC growth hormone deficiency, lowering glucose levels in non-insulin
 CC dependent diabetes mellitus and other disorders involving glucose
 CC intolerance, enhancing mucosal regeneration and as immunosuppressants

XX Sequence 3120 BP; 936 A; 637 C; 706 G; 841 T; 0 U; 0 Other;
 SQ

Alignment Scores:

Pred. No.: 0 Length: 3120
 Score: 4700.00 Matches: 882
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 4 Gaps: 0

US-10-825-632-1 (1-882) x AAC85694 (1-3120)

QY 1 MetAlaAlaMeGluThrGluGlnLeuGlyValGluLeuPheGluThrAlaAspCys 20
 DB 214 ATGGCAGCAGCAATGGAACAGACAGCTGGGTGTGAGATATTGGAACCTGGGACTGT 273
 QY 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
 DB 274 GAGGAGAAATTGGAATACAGGATCGGCTAAATGGAGCCCTTTTATGTGAGCGGAT 333
 QY 41 SerTrpSerGlnLeuLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
 DB 334 TCTTGAGTCAGCTTAAAGAGCTGCTTGGCATACAGAAATATCATGGCTACATGATG 393
 QY 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
 DB 394 GCTAAGGCACCATGATTTTCATGTTGTGAAGAGGAATGATCCAGATGACCTCATTTCA 453
 QY 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
 DB 454 GACAGAAATCTATTACCTTGCCATGCTGTGTGAGAACACAGAGAAATACACTGTTTATCT 513
 QY 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
 DB 514 GAAATTCACAAACATCAATAGACAGCAGCTTAATGCTCTCTTGGAGCCCTCTTTTG 573
 QY 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg 140
 DB 574 GATCTTTTCAGCAACACTGGACTATGGAATGATTTCTCGAGAGAGAACTATTAAAGA 633
 QY 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
 DB 634 GAAAGAAAACCGCATTTGGAACAGTCGGAATTCCTTACGATTATACCAAGGAAGTGA 693
 QY 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
 DB 694 ACATTTCTGTTTCAAGCCGGTAGTGAATTTATCAGGTAAAGATGGAGGGCCACAAAGA 753
 QY 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
 DB 754 TTTACGCAACAACCTTTAAGGCCCAATCTAGTGGAACTAGTTGTCCCAACATACCGATG 813
 QY 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
 DB 814 GATCCAAATTTATGCCCGCTGATCCAGACTGGATTGCTTTTATACATAGCAACGATATT 873

QY 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu 240
 DB 874 TGGATATCTAAACATCGTAACAGAGAGAAAGAGACTCACTTATGTGCAACATGAGCTA 933
 QY 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu 260
 DB 934 GCCAACATCGAAGAGATGCCAGATCAGCTGGAGTCGCTACTCTTTGTTCTCCAAAGAA 993
 QY 261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280
 DB 994 TTTGATAGATATCTGGCTATTGGTGGTGTCCAAAGCTGAACAACACTCCCAAGTGGTGT 1053
 QY 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300
 DB 1054 AAAAATCTTTAGAAATCTATATGAAGAAATGATGAATCTGAGGTGGAAATTTATTCATGTT 1113
 QY 301 ThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThr 320
 DB 1114 ACATCCCTTATGTTGGAACAAGAGGGGCGAGATTCATTCCTGTTATCTTAAACAGGTACA 1173
 QY 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
 DB 1174 GCAATCTCTAAAGTCACITTTTAAAGATGTCAGAAATATGATTTGATGCTGAAGGAAGCATC 1233
 QY 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
 DB 1234 ATAGATGTCATAGATAAGGAACATAATCAACCTTTTGAGATTCCTATTGGAAGGATTTGAA 1293
 QY 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
 DB 1294 TATATTGCCAGAGCTGGATGGACTCTCTGAGGAAATATGCTTGGTCACTACTACTAGAT 1353
 QY 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
 DB 1354 CGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATTTATTTATCCCAAGTAGAA 1413
 QY 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
 DB 1414 GATGATGTTATGGAAGGACAGAGACTCATTTGAGTCAGTCGCTGATTTCTGTGACGCCACTA 1473
 QY 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
 DB 1474 ATTATCTTGAAGAAACAACAGACATCTGGATAATATCCATGACATCTTTTCATGTTT 1533
 QY 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
 DB 1534 CCCAAAGTCACGAGAGGAAATTCAGTTTATTTTGGCTCTGATGCAAAACAGGTTTC 1593
 QY 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
 DB 1594 CGTCATTTATACAAATTTACATCTATTTTAAAGGAAGCAAAATATTAACGATCCAGTGGT 1653
 QY 481 GlyLeuProAlaProSerAspPheLysCysProLysLeuGluGluIleAlaIleThrSer 500
 DB 1654 GGGCTCGCTCCCAAGTCAATTTCAAGTGTCTATCAAAAGAGGAGATAGCAATTTACCAGT 1713
 QY 501 GlyValTrpGluValLeuGluValArgHisGlySerAsnIleGlnValAspGluValArgArg 520
 DB 1714 GGTGAATGGGAAGTCTCTTGCCGGCATGATCTAATATCCAAAGTTGATGAAGTCAGAGG 1773
 QY 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer 540
 DB 1774 CTGATATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTACGTAGTCAGT 1833
 QY 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
 DB 1834 TAGCTAAATCTCTGAGAGGTGACAAGGCTGACTGACCGCTGGCTACTCACAATTTCTTGTGC 1893
 QY 561 IleSerGluHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
 DB 1894 ATCAGTCAGCACTGTGACTCTTTTAAAGTAGTATAGTAAACCAAGAGAAATCCACACTGT 1953

QY 591 ValSerLeuTyrLysLeuSerProGluAspAspProThrCysGlyThrLysGluPhe 600
Db 1954 GTGTCCTTTTCAAGCTATCAAGTCTGAAGATGACCCCACTTGCAAAACAAAGAAATTT 2013
QY 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 620
Db 2014 TGGGCCACCATTTGGATTTCAGCAGGTCCTCTTCCTGACTATACCTCCCAAGAAATTTTC 2073
QY 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
Db 2074 TCTTTTGAAGTACTACTGGATTTACATTGTATGGTACTCTCAAGCCCTCATGATCTA 2133
QY 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 660
Db 2134 CAGCCTGGAAGAAATATCTCTACTGTCTGCTTCATATATGGTGGTCTCAGGTGCAGTTG 2193
QY 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680
Db 2194 GTGAATAATCGTTTAAAGGATCAAGTATTTCCGCTTGAATACCTAGCCTCTCTAGGT 2253
QY 681 TyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 700
Db 2254 TATGTGGTTGTAGTAGACACAGGGGATCTGTCCAGGGGCTTAAATTTGAAAGGC 2313
QY 701 AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr 720
Db 2314 GCCTTTAAATATAAATGGGTCAAAATAGAAATTTGACGATCAGGTGAAGGACTCCAATAT 2373
QY 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyr 740
Db 2374 CTAGCTTCTCGATATGATTTTCATTTGACTTAGATCGTGTGGGCATCCACGGCTGCTTAT 2433
QY 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
Db 2434 GGAGGATACCTCTCCCTGATGCGCATTAATGCAGAGGTTCAGATATCTTCAGGGTTGCTATT 2493
QY 761 AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780
Db 2494 GTTGGGGCCCACTGCTCTGTGATCTTCTATGATACAGGATACACGCACTGTTATATG 2553
QY 781 GlyHisProAspGlnAsnGluGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
Db 2554 GGTACCTCTGACAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAA 2613
QY 801 LysPheProSerGluProAsnArgLeuLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
Db 2614 AAGTTCCCTCTGAAACCAATCGTTTACTCTTACATGTTTCTTGTGATGAGATGTC 2673
QY 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
Db 2674 CATTTTGACATACCAGTATATTTACTGAGTTTTTTAGTGAGGCTGGAAGGCCATATGAT 2733
QY 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
Db 2734 TTACAGATCTATCTCTCAGGAGACACAGCATAAGAGTTCTCTGAATCGGGAGAACATTAT 2793
QY 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLys 880
Db 2794 GAACTGCACTTTTTCACCTACCTTCAAGAAAACCTTGGATCAGGTATTTGCTGCTCAAAA 2853
QY 881 ValIle 882
Db 2854 GTGATA 2859
RESULT 6
ID AAD38956 standard; cDNA; 3120 BP.
XX
AC AAD38956;
XX
XX 23-SEP-2002 (first entry)
XX
DE Human dipeptidyl peptidase 8 (DPP8) cDNA.

XX Human; dipeptidyl peptidase; DPP; neoplasia; type II diabetes; cirrhosis;
KW autoimmunity; human immuno deficiency virus; HIV infection; cytostatic;
KW graft rejection; antidiabetic; antiinflammatory; immunosuppressive;
KW antiviral; enzyme; gene; ss.
XX Homo sapiens.
XX Key Location/Qualifiers
FH 214. .2862
CDS /*tag= a
FT /product= "Human DPP8 protein"
FT
XX WO200234900-A1.
XX PN 02-MAY-2002.
XX PF 29-OCT-2001; 2001WO-AU001388.
XX PR 27-OCT-2000; 2000AU-00001078.
XX (UNSY) UNIV SYDNEY.
XX PA Abbott CA, Gorrell MD;
XX PI WPI: 2002-454646/48.
XX DR P-PSDB; AA224170.
XX New dipeptidyl peptidase (DPP) peptides, useful for screening inhibitors
PT of DPP catalytic activity, which may be employed to treat e.g. neoplasia,
PT type II diabetes, cirrhosis, autoimmunity, graft rejection and HIV
PT infection.
XX Example; Fig 1; 91pp; English.
XX The present invention relates to dipeptidyl peptidase (DPP) proteins and
CC polynucleotides encoding such proteins. The DPP peptides are useful for
CC screening inhibitors of DPP catalytic activity. The inhibitors are useful
CC for treating neoplasia, type II diabetes, cirrhosis, autoimmunity, graft
CC rejection and HIV (human immuno deficiency virus) infection. The present
CC sequence is human DPP8 cDNA
XX
SQ Sequence 3120 BP; 936 A; 637 C; 706 G; 841 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 0 Length: 3120
Score: 4700.00 Matches: 882
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: Gaps: 6
US-10-825-632-1 (1-882) x AAD38956 (1-3120)
QY 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
Db 214 ATGGCAGCAGCATATGAAACAGACAGCTGGGTGTGAGATATTTGAACTGCGACTGT 273
QY 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
Db 274 GAGGAAATATTGAATACAGGATCGGCTAAATGGAGCCCTTTTATGTTGACCGGTAT 333
QY 41 SerTrpSerGlnLeuLysLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
Db 334 TCCTGGAGTCAGCTTAAAAAGCTGCTTCCGATACCAAGAAATATCATGGCTACATGATG 393
QY 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
Db 394 GCTAAGGCACCATGATTTTCATGTTGTGTGAGAGGAATGATCAGATGGACCTCATTTCA 453
QY 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
Db 454 GACAGAATCTATTACCTTCCCATGCTGTGTGAGAAACAGAGAAATATACACTGTTTTTATCT 513

Qy 101 GluIleProLysThrIleAsnArgAlaValLeuMetLeuSerTrpLysProIleuLeu 120
Db 514 GAAATTCACAAACTATCAATAGACGACGCTTAATGCTCTCTTGGAGGCTCTTTTG 573
Qy 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuArg 140
Db 574 GATCTTTTTCAGGCAACACTGGACTATGGAAATGATTTCTCGAAGAAGAACTATTAGA 633
Qy 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
Db 634 GAAAGAAACGCAATTTGGAACAGTCGGAATTCCTTACCAATATCACCAAGAGTGA 693
Qy 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
Db 694 ACAATTTCTGTTTCAAGCGGTAGTGAATTTATCACGTAAGATGGAGGCGGCACAGA 753
Qy 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
Db 754 TTTACGCAACAACTTTAAGGCCCAATCTAGTGGAACTAGTTGTCCCAACATACGATG 813
Qy 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
Db 814 GATCCAAATTAATGCCCCGCTGATCCAGACTGGATTCCTTTATACATAGCAACGATAT 873
Qy 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu 240
Db 874 TGGATATCTCAACATCGTAACAGAGAGAGAGAGAGACTCTATGATGCAACATGAGCTA 933
Qy 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu 260
Db 934 GCCAACATGGGAAGAGATGCCAGATCAGCTGAGTCGCTACCTTTGTTCTCCAAAGAA 993
Qy 261 PheAspArgTyrSerGlyTyrTrpTyrCysProLysAlaGluThrThrProSerGlyGly 280
Db 994 TTTGATAGATATCTTGGCTATTTGGTGGTGTCCAAAGCTGAAACAACTCCAGTGGTGT 1053
Qy 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300
Db 1054 AAAATCTTAGAATTTCTATATGAGAAATGATGAATCTGAGGTGGAAATTTATCATGTT 1113
Qy 301 ThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThr 320
Db 1114 ACATCCCTTATGTTGGAAACAAGAGGAGGAGATTCATTCGGTTATCTCTAAACAGGTACA 1173
Qy 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluArgIle 340
Db 1174 GCAATCTCTAAAGTCACATTTTAAAGATGTCAGAAATATGATGCTGAAGAGGATC 1233
Qy 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
Db 1234 ATAGATGTCATAGATAAGAACTAATTCACCTTTTCAGATTCATTTGAAGAGTTGAA 1293
Qy 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
Db 1294 TATATTCAGAGCTGGATGGACTCTCGAGGAAATATGCTGGTCCATCTCTACTAGAT 1353
Qy 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
Db 1354 CGCTCCCACTCGCTCAGATAGTGTGATCTCACCTGAATTTATTTATCCAGGATGAA 1413
Qy 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
Db 1414 GATGATGTTATGGAAGGAGAGACTCATTTGAGTCAGTGGCTGATCTGTGAGCGCCACTA 1473
Qy 421 IleIleTyrGluGluThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
Db 1474 ATTATCTATGAAGAAACAACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTT 1533
Qy 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
Db 1534 CCCCAGATCAGAAAGAGAAATGAGTTATTTTTCCTCTGAATGCAAAACAGGTTTC 1593

Qy 461 ArgHisLeuTyrIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
Db 1594 CGTCATTTATACAAAATTTACATCTATTTTAAAGGAAAGCAATAATAACGATCCAGTGGT 1653
Qy 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 500
Db 1654 GGGCTGCCTGCTCCCAAGTGATTTCAAGTGTCTATCAAGAGGAGATAGCAATTTACCAGT 1713
Qy 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
Db 1714 GGTGAATGGGAAGTTCTTGGCGGCATGGATCTAATATCCAAAGTTGATGAAGTCAGAAGG 1773
Qy 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValIleSer 540
Db 1774 CTGGTATATTTTGAAGGCACCAAGACTCCCCCTTTAGAGCATCACCTGTAGTAGTCAGT 1833
Qy 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
Db 1834 TAGCTAATCTCTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACATTTCTTGTGTC 1893
Qy 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
Db 1894 ATCAGTCAGCACTGTGACTTCTTTATAAGTAAGTATAGTAAACCAAGAAATCCACACTGT 1953
Qy 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
Db 1954 GTGTCCCTTTTACAAGCTATCAAGTCTCAAGATGCCCACTTCGCAAAACAAGGAATTT 2013
Qy 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 620
Db 2014 TGGGCCACCAATTTTGGATTCAGCAGTCTCTCTCTGACTATATCTCTCCAGAAATTTTC 2073
Qy 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
Db 2074 TCTTTTGAAGTACTACTGGATTTTACATTTGATGGATGCTCTCAAGCCTCATGACTA 2133
Qy 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 660
Db 2134 CAGCCTGGAAAGAAATATCTCTACTGTCTGTTTCATATATGTTGCTCTCAGGTGTCAGTTG 2193
Qy 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680
Db 2194 GTGAATTAATCGGTTTAAAGGAGTCAAGTATTTCCCTTGAATATCCCTAGCCTCTCTAGGT 2253
Qy 681 TyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 700
Db 2254 TATGTGTTGTAGTAGACACAACAGGGGATCTGTCAACGAGGGCTTAAATTTGAAGGC 2313
Qy 701 AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr 720
Db 2314 GCCTTTAAATATAAAATGGGTCAAAATAGAAAATTCAGCATCAGGTGGAGGACTCCAATAT 2373
Qy 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyr 740
Db 2374 CTAGCTTCTCGATATGATTTTCAATGACTTAGATCTGTGGGATCCACGGCTGTCTAT 2433
Qy 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
Db 2434 GGAGGATACCTCTCCCTGATGGCATTAATGCAGAGGTCCAGATATCTTCAGGTTGCTATT 2493
Qy 761 AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780
Db 2494 GCTGGGGCCCCAGTCACCTCTGGGATCTTCTATGATACAGGATACACGGAACGTTATATG 2553
Qy 781 GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
Db 2554 GGTCAACCTGACCAAGTGAACAGGGCTATTACTTAGGATCTGTGGGATGCAAGGAGAA 2613
Qy 801 LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
Db 2614 AGTTCCTCTGGAACCAANTGTTTACTTGTCTTACTGATGTTTCTCTGGATGAGAAATGTC 2673
Qy 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840

DB	2674	GACATACAGTATATACAGAGT GAGGCTGGAAAGCCATATGAT	2733
QY	841	LeuGlnIleTyrProGlnGluA:GHisSerIleArgValProGluSerGlyGluHisTyr	860
DB	2734	TTACAGATCTATCTCTCAGGAGACACAGCATAGAGTTCTCGAATCGGGAGACATPAT	2793
QY	861	GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAalaLeuLys	880
DB	2794	GAACCTGCATCTTTTGCACCTACCTTCAGAAAACCTTTGGATCAGTATGCTCTAATA	2853
QY	881	Vallile 882	
DB	2854	GTGATA 2859	
AAH99934			
ID	AAH99934	standard; cDNA; 3143 BP.	
XX	AC	AAH99934;	
XX	DT	12-APR-2002 (first entry)	
XX	DE	cDNA encoding 21953 human prollyl oligopeptidase.	
XX	KW	21953 prollyl oligopeptidase; human; proline; endopeptidase; cancer;	
XX	KW	cardiovascular disease; autoimmune disease; atopic allergy;	
XX	KW	neural disorder; vascular disorder; prostate disorder; cytostatic;	
XX	KW	diabetic; antiarthritic; antiasthmatic; antiinflammatory;	
XX	KW	diabetes mellitus; arthritis; multiple sclerosis; asthma;	
XX	KW	Grave's disease; neuronal disorder; demyelinating disease; ss.	
XX	OS	Homo sapiens.	
XX	PH	Key	Location/Qualifiers
XX	FT	CDS	229..2877
XX	FT		/*tag= a
XX	FT		/product= "21953 prollyl oligopeptidase"
XX	FT		/note= "This region is specifically claimed in claim 2"
XX	PN	WO200179473-A2.	
XX	PD	25-OCT-2001.	
XX	PF	11-APR-2001; 2001WO-US040483.	
XX	PR	18-APR-2000; 2000US-0197508P.	
XX	PA	(MILL-) MILLENNIUM PHARM INC.	
XX	PI	Meyers RA, Williamson M;	
XX	DR	WPI; 2002-034353/04.	
XX	DR	P-PSDB; AAG78415.	
XX	DR	New polypeptides 21953, member of human prollyl oligopeptidase family,	
XX	PT	useful as diagnostic targets and therapeutic agents for controlling	
XX	PT	cancer, lymphoma and leukemia.	
XX	PT	Claim 7; Page 100-102; 121pp; English.	
XX	PS	This invention relates to an isolated 21953 human prollyl oligopeptidase.	
XX	CC	Which is cytostatic, antidiabetic, antiarthritic, neuroprotective,	
XX	CC	antithyroid, dermatological, antipsoriatic, antiasthmatic,	
XX	CC	ophthalmological, antiinflammatory, nootropic, antiparkinsonian,	
XX	CC	anticonvulsant, gynaecological, vasotropic, antianginal, cardiac,	
XX	CC	antithrombotic, anorectic and metabolic in its action. Uses include	
XX	CC	gene therapy, expression or activity of 21953 protein modulator, it is	
XX	CC	useful for identifying a compound which binds to it and can be used in	
XX	CC	preventing, treating or detecting a cellular proliferative or	
XX	CC	differentiative disorder. The 21953 molecule can act as novel diagnostic	
XX	CC	targets and therapeutic agents for controlling disorders associated with	
XX	CC	the aberrant activity or degradation of peptide hormones e.g., disorders	
CC	CC	immune function, reproductive, neurological and proliferation such as cancer,	
CC	CC	The 21953 molecules are thus useful for treating and preventing cellular	
CC	CC	proliferative and differentiative disorders, haematopoietic neoplastic	
CC	CC	disorders, immune disorders such as autoimmune diseases, diabetes	
CC	CC	mellitus, arthritis, multiple sclerosis, asthma, Grave's disease,	
CC	CC	neural disorders, demyelinating diseases, vascular disorders and	
CC	CC	metabolism or pain disorders. This sequence represents the cDNA encoding	
CC	CC	sequence of 21953 human prollyl oligopeptidase	
XX	XX	Sequence 3143 BP; 943 A; 644 C; 712 G; 844 T; 0 U; 0 Other;	
XX	XX	Alignment Scores:	
XX	XX	Pred. No.: 0 Length: 3143	
XX	XX	Score: 4700.00 Matches: 882	
XX	XX	Percent Similarity: 100.0% Conservatives: 0	
XX	XX	Best Local Similarity: 100.0% Mismatches: 0	
XX	XX	Query Match: 100.0% Indels: 0	
XX	XX	DB: 6 Gaps: 0	
XX	XX	US-10-825-632-1 (1-882) x AAH99934 (1-3143)	
QY	1	MetAlaAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys	20
DB	229	ATGGCAGCAGCAATGGAAACAGACAGCTGGGTGTGAGATATTTGAACTCGCGACTGT	288
QY	21	GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr	40
DB	289	GAGGAGATATTGATATCATCAGATCGGCCTAAATTGGAGCCTTTTATGTGAGCGGTAT	348
QY	41	SerTrpSerGlnLeuLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet	60
DB	349	TCTTGAGTCAGCTTAAAGCTGTTCCTGATACAGAAATATCATGCTACATGATG	408
QY	61	AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer	80
DB	409	GCTAAGGCGACCATGATTTTGTGTGANGAGGAATGATCCAGTAGACCTCATTC	468
QY	81	AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer	100
DB	469	GACAGAAATCTATTACCTTGCCTATGCTGTGTGAGAACAGAGAAATACACCTGTTTATCT	528
QY	101	GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu	120
DB	529	GAATTTCCCAAACTATCAATAGACGACAGCTTAAATGCTCTCTTGGAAGCCTCTTTTG	588
QY	121	AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg	140
DB	589	GATCTTTTTCAGCAACACTGCACTATGGAATGTATTCTCGAGAAGAAGAACTATTAAAG	648
QY	141	GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly	160
DB	649	GAAGGAAAACGATTTGGAAACAGTCGGAATGCTTCTTACGATTATCACCAGGAAGTGA	708
QY	161	ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly	180
DB	709	ACATTTCTGTTTCAAGCCGTAGTGAATTTATCACGTAAAGATGAGAGGCCACAGGA	768
QY	181	PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleAspMet	200
DB	769	TTTACGCAACAACTTTAAGGCCCACTAGTGGAACTAGTTGTCTCCACATACGATG	828
QY	201	AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle	220
DB	829	GATCCAAAAATTTATGCCCTGCTGATCCAGACTGGATTTGCTTTATACATAGCAGATAT	888
QY	221	TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu	240
DB	889	TGGATATCTAAATCATCTTAACACAGAGAAGAAAGAGACTCACTTATGTGCAATGAGCTA	948
QY	241	AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu	260
DB	949	GCCAAACATGGAAGAAGATCCAGATCAGCTGGAGTGTCTTGTCTCTCAAGAGAA	1008

Qy 261 PheAspArgTyrSerGlyTyrTrpCysProLysAlaGluThrThrProSerGlyGly 280
Db 1009 TTTGATAGATATTCCTGGCTATTGGTGTGCTCAAAAGCTGAACCAACTCCAGTGGTGT 1068
Qy 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300
Db 1069 AAAATTCCTTAGAATTCCTATATGAGAAATGATGAATCTGAGGTGGAATATTATCATGTT 1128
Qy 301 ThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThr 320
Db 1129 ACATCCCTATGTTGGAAACAAGAGGAGGAGAGATTCATTCGTTATCTCTAAAACAGGTACA 1188
Qy 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
Db 1189 GCAAACTCTAAAGTCACTTTTAAAGATGTCAGAAATATGATGATGCTGAAGGAAGATC 1248
Qy 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
Db 1249 ATAGATGTCATAGATAAGGAACATAATTCACCTTTTGAGATTTCTATTGGAAGAGTTGAA 1308
Qy 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
Db 1309 TATATTGCCAGAGCTGGATGGACTCTCTGAGGAAATATGCTTGCTCCATCTCTACTAGAT 1368
Qy 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
Db 1369 CGTCTCCAGACTCGCTCGAGATAGTGTGATCTCACCTGAATATTATTTATCCAGTAGAA 1428
Qy 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
Db 1429 GATGATGTTATGGAAGGACAGAGACTCATTTGAGTCAGTGCCTGATTTCTGTGAGCGCCACTA 1488
Qy 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
Db 1489 ATTATCTATGAGAAACAACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTTT 1548
Qy 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
Db 1549 CCCAAAGTCACAGAGAGAAATGATGTTATTTTGGCTCTGATGCAAAACAGGTTTC 1608
Qy 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
Db 1609 CGTCATTTATACAAATTTACATCTATTTTAAAGGAAGCAAAATATAACAGATCCAGTGT 1668
Qy 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 500
Db 1669 GGGCTGCTCTCCAAAGTGAATTTCAAGTGTCTCTATCAAAAGAGAGATAGCAATTTACCAGT 1728
Qy 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
Db 1729 GGTGAATGGGAAGTCTTGGCCGGCATGGATCTAATATCCAAAGTTGATGATGATGAGG 1788
Qy 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer 540
Db 1789 CTGGTATATTTGAGGCCACCAAGACTCCCTTTAGAGCATCACCTGTACGTAGTCAGT 1848
Qy 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
Db 1849 TACGTAATCTCGAGAGAGTGACAAGGCTGACTGACCGTGGCTACTCACATTTCTTGCTGC 1908
Qy 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
Db 1909 ATCAGTCAGCACTGTGACTTCTTTAAGTAGTAGTAGTATACCAAGAGATCCACACTGT 1968
Qy 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
Db 1969 GTGTCCCTTTACAGCTATCAAGTCTGAGATGACCAACTTGGCAAAACAAGAAATTT 2028
Qy 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePhe 620
Db 2029 TGGGCCCACTTTTGGATTGAGGAGTCTCTCTCTGACTATCTCTCTCCAGAAATTTTC 2088

Qy 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
Db 2089 TCTTTTGAAGTACTACTGGATTTTACATGTATGGATGCTCTACAAGCTCATGATCTA 2148
Qy 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 660
Db 2149 CAGCTTGGAAAGAAATATCTCTACTGTGCTGTTTATATATGCTGCTCTCAGGTGCAAGTTG 2208
Qy 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680
Db 2209 GTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTTAGCTCTCTAGGT 2268
Qy 681 TyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 700
Db 2269 TATGTGGTGTGATGATAGACAAACAGGGGATCCTGTCAACGAGGCTTAAATTTGAAGGC 2328
Qy 701 AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr 720
Db 2329 GCCTTTAATATATAAATGGGTCAATAGAAATTCAGCATCAGGTGGAAGGACTCCAATAT 2388
Qy 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyr 740
Db 2389 CTAGCTTCTCGATATGATTTTCACTTAGATCGTGTGGCATCCACGGCTGCTCTAT 2448
Qy 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
Db 2449 GGAAGGATACCTCTCCCTGATGGCATTAATGACAGAGGTCAAGATATCTTTCAGGGTTGCTATT 2508
Qy 761 AlaGlyValAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780
Db 2509 GCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGATACACGGAACGTTATATG 2568
Qy 781 GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
Db 2569 GGTCACTCCCTGACCCAGAAATGACAGGGCTATTACTTAGGATCTGTGGCATCGAAGCAGAA 2628
Qy 801 LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
Db 2629 AAGTCTCCCTCTGAACAAATCGTTTACTGCTCTTACATGTTTCTCTGGATGAGAATGTC 2688
Qy 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
Db 2689 CATTTGGACATACCAAGTATATTACTGAGTTTTTTAGTGAGGGCTGGAAAGCCATATGAT 2748
Qy 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
Db 2749 TTACAGATCTATCTCTCAGAGAGACACAGCATTAAGATTCCTGAATCGGAGAACATTAT 2808
Qy 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLys 880
Db 2809 GAACGTGCATCTTTTGGCACTACCTTCAAGAAAAACCTTGGATCACGTATTGCTGCTCTAAA 2868
Qy 881 ValIle 882
Db 2869 GTGATA 2874
RESULT 8
AAH99935
ID AAH99935 standard; cDNA; 2643 BP.
XX
AC AAH99935;
XX
DT 12-APR-2002 (first entry)
XX
DE Coding sequence of 21953 human prollyl oligopeptidase.
XX 21953 prollyl oligopeptidase; antibody; proline; proline; endopeptidase; cancer;
KW cardiovascular disease; autoimmune disease; atopic allergy;
KW neuronal disorder; vascular disorder; prostate disorder; cytosolic;
KW antidiabetic; antihypertensive; antiaesthetic; antiinflammatory;
KW diabetes mellitus; arthritis; multiple sclerosis; asthma;
KW Grave's disease; neuronal disorder; demyelinating disease; ss.
XX

Tue Apr 18 08:18:32 2006

OS	Homo sapiens.
XX	WO200179473-A2.
PP	25-OCT-2001.
PD	
XX	11-APR-2001; 2001WO-US040483.
XX	18-APR-2000; 2000US-0197508P.
XX	(MILL-) MILLENNIUM PHARM INC.
XX	Meyers RA, Williamson M;
XX	WPI; 2002-034353/04.
XX	P-PSDB; AAG78415.
XX	
XX	New polypeptides 21953, member of human prollyl oligopeptidase family,
PT	useful as diagnostic targets and therapeutic agents for controlling
PT	cancer, lymphoma and leukemia.
XX	
XX	Claim 2; Page; 121pp; English.
PS	This invention relates to an isolated 21953 human prollyl oligopeptidase.
CC	Which is cytostatic, antidiabetic, antiarthritic, neuroprotective,
CC	antithyroid, dermatological, antipsoriatic, nootropic, antiparkinsonian,
CC	ophthalmological, antiinflammatory, vasotropic, antitropical, cardiac,
CC	antiatherosclerotic, anorectic and metabolic in its action. Uses include
CC	gene therapy, expression or activity of 21953 protein modulator. It is
CC	useful for identifying a compound which binds to it and can be used in
CC	preventing, treating or detecting a cellular proliferative or
CC	differentiative disorder. The 21953 molecules can act as novel diagnostic
CC	targets and therapeutic agents for controlling disorders associated with
CC	the aberrant activity or degradation of peptide hormones e.g., disorders
CC	associated with cell differentiation and proliferation such as cancer,
CC	immune function, reproductive, neurological and cardiovascular function.
CC	The 21953 molecules are thus useful for treating and preventing cellular
CC	proliferative and differentiative disorders, haematopoietic neoplastic
CC	disorders, immune disorders such as autoimmune diseases, diabetes
CC	mellitus, arthritis, multiple sclerosis, asthma, Grave's disease,
CC	neural disorders, demyelinating diseases, vascular disorders and
CC	metabolism or pain disorders. This sequence represents the cDNA encoding
CC	sequence of 21953 human prollyl oligopeptidase. This sequence represents
CC	the coding sequence of 21953, being the sequence in between the start and
CC	the stop codon of the sequence represented in AAH99934
XX	
SQ	Sequence 2643 BP; 800 A; 514 C; 585 G; 744 T; 0 U; 0 Other;
Alignment Scores:	
Pred. No.:	0 Length: 2643
Score:	4695.00 Matches: 881
Percent Similarity:	100.0% Conservative: 0
Best Local Similarity:	100.0% Mismatches: 0
Query Match:	99.9% Indels: 0
DB:	6 Gaps: 0
US-10-825-632-1 (1-882) x AAH999935 (1-2643)	
Qy	2 AlaAlaMetGluThrGluGlnLeuGlyValGluLlePheGluThrAlaAspCysGlu 21
Db	1 GCAGCAGCAATGMAACAGACACTGGTGTTCGAGATTATTGAACCTGCAGCTGTGAG 60
Qy	22 GluAsnIleGluSerGlnAspArgProlysLeuGluProPheTyrValGluArgTyrSer 41
Db	61 GAGAATATTGATCATCAGGATCGGCTAAATTTGGAGCCCTTTTATGTGTGAGCGGTATTC 120
Qy	42 TrpSerGluLeuValLeuLeuAlaAspThrArgLySvTyrHisGlyTyrMetMetAla 61
Db	121 TGGAGTCAGCTTAAAAGCTGTTCGCCNATACCAGANAATATCATGCTACATGGCT 180
Qy	62 LysAlaProHisaspPheMetPheValIysArgAsnAspProAspGlyProHisSerAsp 81

CC treatment of reproductive, inflammatory and metabolic disorders. ABK83322
XX -ABK83343 encode human DRP proteins
SQ Sequence 4829 BP; 1466 A; 886 C; 1017 G; 1460 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 0 Length: 4829
Score: 4680.00 Matches: 882
Percent Similarity: 99.8% Conservative: 0
Best local Similarity: 99.8% Mismatches: 0
Query Match: 99.6% Indels: 2
DB: Gaps: 0

US-10-825-632-1 (1-882) x ABK83327 (1-4829)

QY 1 MetAlaAlaMetGluThrGluGlnLeuGluValGluLeuPheGluThrAlaAspCys 20
DB 214 ATGGCAGCAGCAATGGAAACAGAACACAGCTGGGTGTGATATTTGAACATGCCGACTGT 273
QY 21 GluGluAenilleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
DB 274 GAGGAGAAATATTGAATCACAGGATCGGCCTAAATTTGAGCCCTTTTATATTGAGCGGTAT 333
QY 41 SerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
DB 334 TCCTGGAGTCAGCTTAAAGAGCTTGGTTCGCGATACACAGAAATATCATGGCTACATGATG 393
QY 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
DB 394 GCTAAGGCCACCATGATTTTCATGTTTGTCAAGAGGAATGATCCAGATGCACCTCATTTCA 453
QY 81 AspArgLysTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
DB 454 GACGAATCTATTACCTTGCCATGCTCTGGTGAGACAGACAGAAATACACTGTTTTATTCT 513
QY 101 GluLeuProLysThrLysAsnArgAlaValLeuMetLeuSerTrpLysProLeuLeu 120
DB 514 GAAATTCCTCAAACTATCATATAGACAGACGCTTAATGCTCTCTTGGAGCTCTTTTG 573
QY 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140
DB 574 GATCTTTTTCAGGCACACTGGACTATGGAATGATTTCTCGAAGAGAGACTATTNAGA 633
QY 141 GluArgLysArgLysGlyThrValGlyLeuAlaSerTyrAspTyrHisGlnGlySerGly 160
DB 634 GAAAGAAAACGCAATTTGGAAACAGTCGGAATTCCTTACCATATATCACCAGAGAGTGA 693
QY 161 ThrPheLeuPheGlnAlaGlySerGlyLysTyrHisValLysAspGlyGlyProGlnGly 180
DB 694 ACATTTCTGTTTCAAGCCGGTAGTGGAAATTTATCACGTAAAGATGGAGGCCACAGGA 753
QY 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnLysArgMet 200
DB 754 TTTACGCAACAACCTTTAAGGCCCAATCTAGTGGAAACTAGTTGTCCCAACATACGGATG 813
QY 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheLysHisSerAsnAspIle 220
DB 814 GATCCAAATATTGCTCTGCTATCCAGACTGGATTGCTTTTATACATACCAACGATATT 873
QY 221 TrpIleSerAsnIleValThrArgGluGluArgArgLeuThrTyrValHisAsnGluLeu 240
DB 874 TGGATATCTAAACATCGTACACAGAGAGAGAGAGAGACTCACTTATGTGCAACATGAGCTA 933
QY 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu 260
DB 934 GCCACATATGGAGAGAGATGCCAGATCAGCTGGAGTGCCTGCTTTGCTTCCCAAGAGAA 993
QY 261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280
DB 994 TTTGATAGATATTCTGGCTATTTGGTGTCTCCAAAGCTGAAACACTCCAGTGGTGGT 1053
QY 281 LysIleLeuArgLysLeuTyrGluLysAsnAspGluSerGluValGluLysIleHisVal 300

DB 1054 AAAATCTTATAGAAATCTTATATGAGAGAAATGATGAAATCTGAGGTGGAAATATTATCATGTT 1113
QY 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320
DB 1114 ACATCCCTCTATGTTGGAAACAAGAGGCGAGATTATCCGTATATCTTAAACAGGTACA 1173
QY 321 AlaAsnProLysValThrPheLysMetSerGluLeuMetIleAspAlaGluGlyArgIle 340
DB 1174 GCAAAATCTTAAAGTCATCTTTTAAAGATGTGAGAAATAATGATGCTGAGAGGAAGATC 1233
QY 341 IleAspValIleAspLysGluLeuGlnProPheGluLeuLeuPheGluGlyValGlu 360
DB 1234 ATAGATGTCTAGATAGAGAACTAAATCAACCTTTTGAAGATTTCTATTGAGAGGATTGAA 1293
QY 361 TyrIleAlaArgAlaGlyTyrThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
DB 1294 TATATTGCCAGAGCTGGATGCACTCTCGAGGAAAAATATGCTGGTGCTCATCTCTACTAGAT 1353
QY 381 ArgSerGlnThrArgLeuGlnIleValLeuLeuIleSerProGluLeuPheIleProValGlu 400
DB 1354 CGCTCCAGACTCGCCTACAGATAGTGTGGATCTCACCTGAAATTTATTTATCCAGTAGAA 1413
QY 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
DB 1414 GATGATGTTATGGAAGAGAGAGACTTCATTTGAGTCTGAGTCTGATTTCTGTGACGCCACTA 1473
QY 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
DB 1474 ATTATCTATGAAGAAACACACAGCATCTGGATAAATATCCATGACATCTTTTCATGTTTT 1533
QY 441 ProGlnSerHisGluGluGluLeuGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
DB 1534 CCCCAAGTCAAGAGAGAAATTTGAGTTATTTTGGCTTGGATGCAAAACAGGTTTC 1593
QY 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
DB 1594 CGTCATTTTATCAAAATTTACATCTATTATTAAGGAAGAAACAAATATTAACATCCAGTGT 1653
QY 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluLeuLysIleAlaIleThrSer 500
DB 1654 GGGCTCGCTCTCCAAAGTGAATTTCAAGTGTCTTATCAAGAGGAGGATAGCAATTCAGT 1713
QY 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
DB 1714 GGTGAATGGAAAGTTCTTTGGCGCGCATGGATCTTAATATCCAAGTTGATGAAGTCAGAAG 1773
QY 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer 540
DB 1774 CTGGTATATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACCTGTACGATGTCAGT 1833
QY 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
DB 1834 TAGTAAATCTCTGGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACATTTCTTGTCTC 1893
QY 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
DB 1894 ATCAGTCAGACCTGTGACTCTCTTTATTAAGTAAATGATTAACCAAGAAATCCACACTGT 1953
QY 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
DB 1954 GTGTCCCTTTACAGCTATCAAGTCTCTGAAGATGACCAACTTGCAGAAACAAAGGAATTT 2013
QY 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 620
DB 2014 TGGGCCACCACTTTTGGATTCAGAGTCTCTCTGACTATATCTCTCCAGAAATTTTC 2073
QY 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
DB 2074 TCTTTTGAAGTAGTACTACTGATTTTACATTTGATGGATGCTCTACAGCCTCATGACTA 2133
QY 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGly--ProGlnValGln 660
DB 2134 CAGCTCGAAAGAAATATCTCTACTGTGCTGTTTCATATATATGGTGGTCTCTCAGGTGCACT 2193

QY 660 euValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAenThrLeuAlaSerLeuG 680
DB 2194 TGGTGAATAATCGGTTTAAAGGAGTCAAGTATTTCGGCTTGAATACCTAGCTCTCTAG 2253
QY 680 lYTyRValValValValLysAsnArgGlySerCysHisArgGlyLeuLysPheGluG 700
DB 2254 GTTATGTGTGTGTAGTAGACACACAGGGGATCTGTGTACACGAGGGCTTAAATTTGAAG 2313
QY 700 lYAlaPheLysTyrLysMetGlyGlnLleGluLleAspAspGlnValGluGlyLeuGlnT 720
DB 2314 GGCCTTTAAATATATAATATGGTCAATAGAAATTCAGATCAGGTGGAGGACTCCAT 2373
QY 720 YrLeuAlaSerArgTyrAspPheLleAspLeuAspArgValGlyLleHisGlyTyrSerT 740
DB 2374 ATCTAGCTTCTCGATATGATTTCAATTCATTCAGTTCGTGTGGCATCCACGGCTGTCTCT 2433
QY 740 YrGlyGlyTyrLeuSerLeuMetAlaLleMetGlnArgSerAspLlePheArgValAlaI 760
DB 2434 ATGGAGGATACCTCTCCCTGATGTCATTAATGTCAGAGGTTCAGATATCTTCAGGGTTGCTA 2493
QY 760 lAlaGlyAlaProValThrLeuTyrPheTyrAspThrGlyTyrThrGluArgTyrM 780
DB 2494 TTGCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGATACAGGACGTTATA 2553
QY 780 etGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaG 800
DB 2554 TGGGTCACTGACCCAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAG 2613
QY 800 lYlAsPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnV 820
DB 2614 AAAAGTTCCCTCTGAACCAATCGTTTACTGTCTTTACATGTTTCTCGGTGAGAAATG 2673
QY 820 aHisPheAlaHisThrSerLleLeuLeuSerPheLeuValArgAlaGlyLysProTyrA 840
DB 2674 TCCATTTTGCACATACCAAGTATATTTACTGAGTTTTTTAGTGAGGGCTGGAAAGCCATATG 2733
QY 840 sPLeuGlnLleTyrProGlnGluArgHisSerLleArgValProGluSerGlyGluHisT 860
DB 2734 ATTTACAGATCTATCTCTCAGGAGACACACAGATAGAGTTCTGTAATCGGAGAACATT 2793
QY 860 YrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgLleAlaLeuL 880
DB 2794 ATGAATGCACTCTTTTGCACTACCTTCAAGAAAACCTTGGATCACGTTATGTGCTCTAA 2853
QY 880 ysValIle 882
DB 2854 AAGTGATA 2861
RESULT 10
ADL13374
ID ADL13374 standard; cDNA; 2797 BP.
AC AC
XX ADL13374;
XX
DT 06-MAY-2004 (first entry)
XX
DE Human steroid-induced C3A liver cell cDNA #1103.
XX
KW ss; gene; Hepatotropic; Gene therapy; Wilson disease; liver disorder;
KW steroid therapy; cirrhosis; hepatitis; human; C3A liver cell.
XX
OS Homo sapiens.
XX
XX US6673549-B1.
PN
XX
PD 06-JAN-2004.
XX
XX 12-OCT-2001; 2001US-00976594.
PF
XX
PR 12-OCT-2000; 2000US-0240409P.
XX
XX (INCY-) INCYTE CORP.

XX Furness LM, Buchbinder JL;
PI
XX WPI; 2004-068610/07.
DR
XX
PT Combination useful for preparing a composition for treating liver
PT disorders associated with steroid therapy, e.g., cirrhosis or hepatitis,
PT comprises cDNAs that are differentially expressed in response to steroid
PT treatment.
XX
PS Claim 1; SEQ ID NO 1103; 141pp; English.
XX
CC The invention relates to a combination comprising cDNAs that are
CC differentially expressed in response to steroid treatment. Also included
CC are the following: a high throughput method for using a cDNA to detect
CC differential expression of nucleic acids in a sample; and a high
CC throughput method of screening molecules or compounds to identify a
CC ligand that specifically binds a cDNA. The sample is from a subject with
CC Wilson disease and comparison of a standard defines a stage of that
CC disease. The high throughput method of screening molecules or compounds
CC to identify a ligand that specifically binds a cDNA comprises: combining
CC the combination with molecules or compounds under conditions to allow
CC specific binding; and detecting specific binding between each cDNA and at
CC least one molecule or compound. The molecules or compounds are regulatory
CC proteins. The combination is useful for preparing a composition for
CC treating liver disorders associated with steroid therapy, e.g., cirrhosis
CC or hepatitis. The present sequence represents a human cDNA which is
CC differentially expressed in steroid-induced C3A liver cells. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format directly from USPTO
CC at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 2797 BP; 866 A; 551 C; 595 G; 785 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 0 Length: 2797
Score: 4443.50 Matches: 840
Percent Similarity: 99.5% Conservative: 0
Best Local Similarity: 99.5% Mismatches: 1
Query Match: 94.5% Indels: 3
DB: 12 Gaps: 1
US-10-825-632-1 (1-882) x ADL13374 (1-2797)
QY 42 TrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMetAla 61
DB 3 TGGAGTCAGCTTAAAGAGCTGCTGCCGATACCAGAAATATCATGCTACATGATGCT 62
QY 62 LysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSerAsp 81
DB 63 AAGGCACCAATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTCAGAC 122
QY 82 ArgLleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAenThrLeuPheTyrSerGlu 101
DB 123 AGAATCTATTACCTTGCCATGCTGTGTGAGAACAGAGAAATATACATGTTTTATCTGAA 182
QY 102 IleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTyrLysProLeuLeuAsp 121
DB 183 ATTCCCAAAACTCATATAGACAGCAGCTCTTAATGCTCTCTTGGAGGCTCTTTTGGAT 242
QY 122 LeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArgGlu 141
DB 243 CTTTTTCAGCAACACTGGACTATGGAATGTATTCTCGAGAAGAAGAACTATTAAAGAGAA 302
QY 142 ArgLysArgLleGlyThrValGlyLleAlaSerTyrAspTyrHisGlnGlySerGlyThr 161
DB 303 AGAAGACGCAATTTGGACAGTCGGAATTCCTTACGATTATACCAAGGAAGTGGAGCA 362
QY 162 PheLeuPheGlnAlaGlySerGlyLleTyrHisValLysAspGlyGlyProGlnGlyPhe 181
DB 363 TTTCTGTTTCAGCCGGTAGTGAATTTATCATCGTAAAGATGGAGGCCACACAGGATTT 422
QY 182 ThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMetAsp 201

423	Db		ACGCAACAACTTTTAAGGCCCAATCTAGTGGAAACCTAGTTGTGCCCAACATACGGATGGAT	482
202	Qy		ProLysLeuCyseProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIleTrp	221
483	Db		CCAAATATATGCGCTGCTGATCCAGACTGGATGCTTTTATACATAGCAACGATATTGG	542
222	Qy		IleSerAsnIleValThrArgGluGluArgLeuThrTyValHisAsnGluLeuAla	241
543	Db		ATATCTAACATCTACACAGAGAAAGAGAGACTCATTATGTGCACAAATGAGTAGCC	602
242	Qy		AsnMetGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGluGluPhe	261
603	Db		AACATGGAAGAGATCCAGATCAGCTGGATCGCTACTCTTTGTTCTCCAGAGAAATTT	662
262	Qy		AspArgTySerGlyTyTrpTrpCyseProLysAlaGluThrThrProSerGlyGlyLys	281
663	Db		GATAGATATCTCGCTATTTGGTGTCCAAAGCTGGAACAACTCCCAAGTGGTGTA	722
282	Qy		IleLeuArgIleLeuTyGluGluAenAspGluSerGluValGluIleIleHisValThr	301
723	Db		ATTCCTTGAATCTTATATGAAGAAATGATGAATCTGAGTGGAAATATTATCATGTACA	782
302	Qy		SerProMetLeuGluThrArg-ArgAlaAspSerPheArgTyTrpLysThrGlyThrAl	321
783	Db		TCCCTATGTGGAACAAAGCGAGGCAGATTCATTCCGTATATCTTAAACAGGTPACGC	842
321	Qy		aAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIleI	341
843	Db		AAATCTTAAAGTCACCTTTTAAAGTGCAGAAATAATGATGTGCTGAAGGAAGATCAT	902
341	Qy		aAspValIleAspLysGluLeuLeuGlnProPheGluIleLeuPheGluGlyValGluTy	361
903	Db		AGATGTCATAGATAAGGNACTAATTCACCTTTTGAGATTCCTATTGGAAGGAGTTGAATA	962
361	Qy		rIleAlaArgAlaGlyTrpThrProGluGlyLys---TyrAlaTrpSerIleLeuLeuA	380
963	Db		TATTCACAGAGCTGGATGGACTCTCGAGGMAAATAGTCTGCTGGTCCATCTACTAG	1022
380	Qy		spArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValG	400
1023	Db		ATCGCTCCAGACTCGGCTACAGATAGTGTGATCTACCTGAATATTATTATCCAGTAG	1082
400	Qy		luAspAspValMetGluArgGluArgLeuIleGluSerValProAspSerValThrProL	420
1083	Db		AAGATGATGTTATGGAAGGACAGACTCATGATGAGTGTGATCTACCTGAATATTATGACGCCAC	1142
420	Qy		euIleIleTyGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValP	440
1143	Db		TAATATATCTATGAAGAAACACACATCTGGATAAATATCCATGACATCTTTCAGTTT	1202
440	Qy		heProGlnSerHisGluGluGluPheIlePheAlaSerGluCyLysThrGlyP	460
1203	Db		TTCGCCAAAGTCAGCAAGAGAGAAATTAAGTTTATTTTTCCTCTGAATCAAAAAGGTT	1262
460	Qy		heArgHisLeuTyTyLysThrSerIleLeuLysGluSerLysTyTrpLysArgSerSerG	480
1263	Db		TCCGTCTATTTATACAAATATACATCTATTTTAAAGGAAAGCAAAATATAAACCATCCAGTG	1322
480	Qy		lyGlyLeuProAlaProSerAspPheLysCyseProIleLysGluGluIleAlaIleThrS	500
1323	Db		GTGGGCTGCTGCTCCAAAGTATTTCAAGTGTCTATCAAGAGAGAGATAGCAATATACCA	1382
500	Qy		erGlyGluTrpGluValLeuGlyVargHisGlySerAsnIleGlnValAspGluValArgA	520
1383	Db		GTGGTGAATGGAAATCTTTGGCGGCATCGATCTAATATCCAAAGTTGATGAAGTCAGAA	1442
520	Qy		rgLeuValTyPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyTyValValS	540
1443	Db		GGCTGGTATATTTTGAAGGCACCAAGACTCCCTCTTTAGAGCATCACCTGTAGTAGTCA	1502
540	Qy		erTyValAsnProGlyGluValThrArgLeuThrAspArgGlyTySerHisSerCysC	560

RESULT 11
ACA92421
ID ACA9

XX AC ACA92421;
XX DT 15-JUL-2003 (first entry)
XX DE DNA encoding human PMMM-6.
XX
KW Human; protein modification and maintenance molecule; PMMM; cancer;
KW cell proliferation disorder; atherosclerosis; neurological disorder;
KW epilepsy; Huntington's disease; stroke; immune disorder; allergy;
KW inflammatory disorder; AIDS; developmental disorder; hypothyroidism;
KW Cushing's syndrome; gastrointestinal disorder; epithelial disorder;
KW infection; cycostatic; antiarteriosclerotic; anticonvulsant; neuroptic;
KW neuroprotective; cerebroprotective; anti-HIV; antiallergic; vulnerary;
KW antiinflammatory; thyromimetic; gene; ds.
XX OS Homo sapiens.
XX PN WO2003031939-A2.
XX PD 17-APR-2003.
XX PF 11-OCT-2002; 2002WO-US032850.
XX PR 12-OCT-2001; 2001US-0329689P.
XX PR 25-OCT-2001; 2001US-0335703P.
XX PR 09-NOV-2001; 2001US-0348887P.
XX PR 28-NOV-2001; 2001US-0334145P.
XX PR 06-DEC-2001; 2001US-0337451P.
XX PR 14-DEC-2001; 2001US-0340584P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Ramkumar J, Gorvad AE, Baughn WR, Emerling BM, Yang J, Lee SY;
XX Tran UK, Becha SD, Duggan BW, Lee EA, Griffin JA, Li JX;
XX Sprague JW, Hafalia AJA, Chawla NK, Lehr-Mason PM, Kable AE, Yue H;
XX Marquis JP, Yao MG, Richardson TW, Tang TY, Jin P, Chien D;
XX Bhatia U, Burrill JD, Lee S, Blake JJ, Ho A, Zheng W;
XX
XX WPI; 2003-430274/40.
XX P-PSDB; ABU92026.
XX
XX New human protein modification and maintenance molecules (PMMM), useful
XX for diagnosing, treating and preventing diseases or conditions associated
XX with the aberrant PMMM expression e.g. cancer, atherosclerosis, or
XX infections.
XX
XX Claim 5; Page 285-286; 31pp; English.
XX
XX The present invention relates to the isolation of human protein
XX modification and maintenance molecules (PMMM), and the polynucleotide
XX sequences encoding them. A total of 40 PMMM polypeptides (designated PMMM
XX -1 to PMMM-40) are disclosed. The sequences of the invention are useful
XX for diagnosing a condition or disease associated with the expression of
XX PMMM in a subject, preparing a polyclonal or monoclonal antibody, and
XX generating an expression profile of a sample containing the
XX polynucleotides. The diseases or conditions associated with decreased
XX expression or overexpression of PMMM are cell proliferation disorders
XX (e.g. cancer, atherosclerosis), neurological disorders (e.g. epilepsy,
XX Huntington's disease, stroke), immune/inflammatory disorders, (e.g. AIDS,
XX allergies), developmental disorders (e.g. hypothyroidism, Cushing's
XX syndrome), gastrointestinal or epithelial disorders, and infections. The
XX PMMM polypeptides or their fragments are useful in screening compounds
XX for effectiveness as agonists or antagonists of the polypeptides, or in
XX altering the expression of the target polynucleotide and compounds that
XX specifically bind to, or modulate the activity of the polypeptide.
XX ACA92416-ACA92455 encode the human PMMM polypeptides of the invention
XX
XX Sequence 2952 BP; 886 A; 604 C; 664 G; 798 T; 0 U; 0 Other;
SQ

Alignment Scores:
Pred. No.: 0 Length: 2952
Score: 4397.50 Matches: 831

Percent Similarity: 94.2% Conservative: 0
Best Local Similarity: 94.2% Mismatches: 51
Query Match: 93.6% Indels: 51
DB: 10 Gaps: 1
US-10-825-632-1 (1-882) x ACA92421 (1-2952)
Qy 1 MetAlaAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
Db 204 ATGGCAGCAGCAATGGAACAGAACTGGGTGTGAGATATTTGAAACTGGGACTGT 263
Qy 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
Db 264 GAGAGAAATTTGAATTCACAGATCGGCTAAATTTGAGCCCTTTTATGTGAGCGTAT 323
Qy 41 SerTrpSerGlnLeuLysLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
Db 324 TCCTGGAGTCAGCTTAAAGAGCTGCTGCCGATACCAGAAATATCATGGCTACATGATG 383
Qy 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
Db 384 GCTAAGGCACCATGATTTTCATGTTTGTGAAGAGGATGATCCAGATGGACCTCATTTCA 443
Qy 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
Db 444 GACAGAAATCTATTACCTTGCCATGTCTGGTGAAGACAGAGAAATACACTGTTTATCT 503
Qy 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
Db 504 GAAATTTCCCAAACTACTCAATAGACAGACAGTCTTAATGCTCTCTTGAAGCCCTTTTG 563
Qy 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg 140
Db 564 GATCTTTTTCAGGCAACACTGGACTATGGAATGATTTCTCGAGAAGAAGAACTATTAA 623
Qy 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
Db 624 GAAAGAAAAACGCAATTTGGAACAGTCGGAATTTGCTTACGATTTATCCCAAGGAAGTGA 683
Qy 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
Db 684 ACATTTCTGTTTCAGCGGTAGTGGATTTATCACGTAAAGATGGAGGCCACCAAGGA 743
Qy 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
Db 744 TTTACGCAACACCTTTAAGGCCCAATCTAGTGGAAACTAGTTGTCCCAACATACGGATG 803
Qy 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
Db 804 GATCCAAAATTTATGCCCTGCTGATCCAGACTGGATGCTTTTATACATAGCAACGATTT 863
Qy 221 TrpIleSerAsnIleValThrArgGluGluArgArgLeuThrTyrValHisAsnGluLeu 240
Db 864 TGGATATCTAACATCTGTAAACAGAGAAAGAGAGACTCACTTATGTGCACATAGAGCTA 923
Qy 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu 260
Db 924 GCCAAATCGAAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAGAGAA 983
Qy 261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280
Db 984 TTTGATAGATATTTCTGGCTATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1043
Qy 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHisVal 300
Db 1044 AATATCTTAGAATCTTATATGAAGAAATGATGAATCTGAGGTGGAATTTATTCATGTT 1103
Qy 301 ThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThr 320
Db 1104 ACATCCCTATGTTGGAACAGAGGGGCGAGATTCATTCCTCGTTATCTTAAACAGGTACA 1163
Qy 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340

Db	1164	CAAATCTCTAAAGTCACCTTTTAAGATGTCAGAAAATTAATGATTGATGCTGAAGGAGGATC	1223
Qy	341	IleAspValIleAspLysGluLeuIleGlnProPheGluIlePheGluGluValGlu	360
Db	1224	ATAGATGTCATAGATAAGGAACCTAATATCAACCTTTTGAGATCTATTTGAAGGAGTTGAA	1283
Qy	361	TyrIleAlaArgAlaGlyThrProGluGlyLysThrAlaThrSerIleLeuLeuAsp	380
Db	1284	TATATTGCCAGAGCTGATGGACTCCTGAGGGGAAATATGCTGGTGCATCTCTAGAT	1343
Qy	381	ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu	400
Db	1344	CGCTCCACAGACTCGCCCTACAGATAGTGTGATCTCACCTGAAATATTTATCCCAAGTAGAA	1403
Qy	401	AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu	420
Db	1404	GATGATGTTATGGAAAGCGCAGAGACTCAITGAGTCAGTGCCTGATTCGTGACGCCACTA	1463
Qy	421	IleIleTyrGluGluThrThrAspIleThrIleAsnIleHisAspIlePheHisValPhe	440
Db	1464	ATTATCTATGAAGAAACAACAGACATCTGGATAAATATCATGACATCTTTTCATGTTTTT	1523
Qy	441	ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe	460
Db	1524	CCCCAAGTCACGAAGAGGAAATTGAGTTATTTTTTGCCTCTGAATGCCAAACACAGTTTC	1583
Qy	461	ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly	480
Db	1584	CGTCATTTATCAAAATTTACATCTATTTTTTAAGAAAGCAATATATACGATCCAGTGGT	1643
Qy	481	GlyLeuProAlaProSerAspPheLysCysProIleLysGluIleAlaIleThrSer	500
Db	1644	GGGCTGCTGCTCCAAAGTGATTTCAAGTGCTCTATCAAGAGAGATAGCAATACCAAGT	1703
Qy	501	GlyGluThrGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg	520
Db	1704	CGTGAATGGGAAGTCTTGGCCGCGCATGGATCTAAATATCAAGTGTGATGAAGTCAGAAG	1763
Qy	521	LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer	540
Db	1764	CTGGTATATTTTGAAGGCACCAAGAGACTCCCTTTTGAAGCATCACCTGTACGTACGT	1823
Qy	541	TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys	560
Db	1824	TACGTAAATCTCGAGAGAGGTGACAAAGCTGACTGACGCTGGCTACTCACATCTCTTGCTGC	1883
Qy	561	IleSerGlnHisCysAspPheIleSerLysTyrSerAsnGlnLysAsnProHisCys	580
Db	1884	ATCAGTCAGCACTGTGACTCTCTTTAAGTAAGTATAGTAAACCAGAGAATCCACACTGT	1943
Qy	581	ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe	600
Db	1944	GTGTCCTTTACAGCTATCAAGTCTCTGAAGATGACCCCACTTGCAAAACAAGGAATTT	2003
Qy	601	ThrAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe	620
Db	2004	TGGGCACATTTTGGATTACAGAGGTCCTCTCTCGACTATCTCTCCAGAAATTTTC	2063
Qy	621	SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu	640
Db	2064	TCTTTTGAAGTACTACTGGATTTTACATTTGATGGGATGCTCTACAAAGCCTCATGATCTA	2123
Qy	641	GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu	660
Db	2124	CAGCCTGGAAAGAAATATCTCTACTGTCTGTTCATATATGTTGGTGGTCTCAGGTGAGTTG	2183
Qy	661	ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly	680
Db	2184	GTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCCCTAGCCTCTCTAGT	2243
Qy	681	TyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly	700
Db	2244	TATGTGGTTGATGATAGACAAAGGGGATCTCTGTCCAGGGGCTTAAATTTGAAGGC	2303
Qy	701	AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr	720
Db	2304	GCCTTTAAATATAAATG-----	2321
Qy	721	LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrPheTyr	740
Db	2321	-----	2321
Qy	741	GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle	760
Db	2322	-----GTTGCTATT	2330
Qy	761	AlaGlyAlaProValThrLeuThrIlePheTyrAspThrGlyTyrThrGluArgTyrMet	780
Db	2331	GCTGGGGCCCGACGTCACCTCTGTGGATCTCTATGATACAGATACACGGAACGTTATATG	2390
Qy	781	GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu	800
Db	2391	GGTCACCTCGACCAAAATGAACAGGGCTATTACTTAGGATCTCTGGCCATGCAAGCAGAA	2450
Qy	801	LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal	820
Db	2451	AAAGTTCCCTCTGAACCAAAATCGTTTACTGCTCTTACATGTTCTCTGGATGAGAATGTC	2510
Qy	821	HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp	840
Db	2511	CATTTTGCACATACCAAGTATATTACTGAGTTTTTTAGTAGGGCTGGAAGCCATATGAT	2570
Qy	841	LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr	860
Db	2571	TTACAGATCTATCTCAGGAGACACAGCATTAAGAGTCTCTGAATCTGGAGAACATTAT	2630
Qy	861	GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLys	880
Db	2631	GAAGTCAATCTTTGCACTACTCTCAAGAAAACCTTGGATCAGCTATTGCTGCTCTAAA	2690
Qy	881	ValIle 882	
Db	2691	GTGATA 2696	
RESULT 12			
ID	ABK83332	standard; cDNA; 4685 BP.	
XX	ABK83332;		
XX	12-AUG-2002 (first entry)		
XX	cdNA encoding human DPRP-1 splice variant #8.		
XX	Human; serine protease; dipeptidyl peptidase IV-related protein; DPRP;		
XX	DPP4; infection; human immunodeficiency virus; HIV-1; HIV-2; pain;		
XX	diabetes; infertility; obesity; anorexia; Parkinson's disease; stroke;		
XX	heart failure; hypertension; urinary retention; osteoporosis; cancer;		
XX	ulcer; allergy; cancer; psychotic disorder; neurological disorder;		
XX	dyskinesia; reproductive disorder; inflammatory disorder;		
XX	metabolic disorder; gene; ss.		
XX	Homo sapiens.		
XX	WO200231134-A2.		
XX	18-APR-2002.		
XX	12-OCT-2001; 2001WO-US031874.		
XX	12-OCT-2000; 2000US-0240117P.		
XX	(PERR) FERRING BV.		
XX	Qi S, Akineanya KO, Riviere PJ, Junien J;		

DR WPI; 2002-444178/47.
 XX P-PSDB; ABG61601.
 PT New dipeptidyl peptidase IV-related proteins and nucleic acids encoding
 PT the proteins, useful for treating e.g. fungal, bacterial, protozoan and
 PT viral infections, cancers, allergies, neurological disorders, or pain.
 XX
 PS Disclosure; Page 75-76; 113pp; English.
 XX
 CC The present invention relates to the isolation of novel human serine
 CC proteases referred to as dipeptidyl peptidase IV (DPPIV)-related proteins
 CC (DPRP). The dipeptidyl peptidase IV-related proteins (DPRP) and nucleic
 CC acids encoding them are useful for treating infections such as fungal,
 CC bacterial, protozoan and viral infections, particularly infections caused
 CC by human immunodeficiency virus (HIV-1 or HIV-2), pain, diabetes,
 CC precocious puberty, infertility, obesity, anorexia, bulimia, Parkinson's
 CC disease, acute heart failure, hypotension, hypertension, urinary
 CC retention, osteoporosis, angina pectoris, stroke, ulcers, asthma,
 CC allergies, cancers, migraine, vomiting, psychotic and neurological
 CC disorders (e.g. anxiety, dementia, or schizophrenia), and dyskinesias.
 CC These may also be used in discovering therapeutic agents for the
 CC treatment of reproductive, inflammatory and metabolic disorders. ABK83322
 CC -ABK83343 encode human DPRP proteins
 XX
 SQ Sequence 4685 BP; 1430 A; 853 C; 991 G; 1411 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 0 Length: 4685
 Score: 4395.50 Matches: 834
 Percent Similarity: 94.6% Conservative: 0
 Best Local Similarity: 94.6% Mismatches: 1
 Query Match: 93.3% Indels: 48
 DB: 6 Gaps: 1

US-10-825-632-1 (1-882) x ABK83332 (1-4685)

Qy 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
 Db 214 ATGCAGCAGCAATGGAACAGACAGCTGGGTGTGAGATATTGGAACCTGGGACTGT 273
 Qy 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
 Db 274 GAGGAGAAATATTGAATCACAGGATCGGCTAAATTTGGAGCCCTTTTATGTTGAGCGGTAT 333
 Qy 41 SerThrSerGlnLeuLysIleLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
 Db 334 TCTGGAGTCAGCTTAAAGCTGCTTGGCAGATACCAAGAAATATCATGCTACATGATG 393
 Qy 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
 Db 394 GCTAAGGCACCAATGATTTTCATGTTTGTGAAGAGATGATCCAGATGGACCTCATTC 453
 Qy 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
 Db 454 GACAGAATCTATTACCTTGCATGCTGCTGAGAACAGAGAAATAACACTGTTTATTCT 513
 Qy 101 GluIleProLysThrIleAsnArgAlaValLeuMetLeuSerTrpLysProLeuLeu 120
 Db 514 GAAATTCCTCCAAATATCAATAGACAGCAGCTTTAATGCTCTCTTGAAGCCCTCTTTG 573
 Qy 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg 140
 Db 574 GAUCTTTTTCAGCAACACCTGGACTATGGAATGATTTCTCGAAGAGAAAGAACTATTAGA 633
 Qy 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
 Db 634 GAAAGAAAAGCGCATGGACAGTCGGAATGCTCTTACGATTATCATCCAAAGGAAGTGA 693
 Qy 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
 Db 694 ACATTTCTGTTCAAGCCGTTAGTGAATTTATCATCGTAAAGATGGAGGGCCACAGGA 753
 Qy 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200

Db 754 TTTTCCCAACAACTTTTAAGGCCCACTAGTGGAACTAGTTGTCCTCCACATACCGATG 813
 Qy 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
 Db 814 GATCCAAATATTATGCTGCTGATCCAGACTGGATGCTTTTATATACATACCAACGATATT 873
 Qy 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu 240
 Db 874 TGGATATCTAACATCGTAACCCAGAGAGAAAGAGACTCACTTATGTGCAATAGACTA 933
 Qy 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu 260
 Db 934 GCCAATGGAAGAGATGCCAGATCAGCTGGTGGCTACCTTTGTTTCTCCAGAGAGAA 993
 Qy 261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280
 Db 994 TTTGATAGATATTCTGCTATTGGTGGTGTCCAAAGCTGAACAACTCCCAAGTGGTGT 1053
 Qy 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300
 Db 1054 AAAATTTCTTAGAATTTCTATATGAGAAAATGATGAATCTGAGGTGGAAATTTATTCATGTT 1113
 Qy 301 ThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThr 320
 Db 1114 ACATCCCTTATGTTGGAACAAGGAGGCGAGATTCATTCCTGTTATCTTAAACAGGTACA 1173
 Qy 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
 Db 1174 GCAATCTCTAAGTCACTTTTAAGATGTCAGAAATATGATGATGCTGAAGGAAGGATC 1233
 Qy 341 IleAspValIleAspLysGluLeuGlnProPheGluIleLeuPheGluGlyValGlu 360
 Db 1234 ATAGATGTCATAGATAAGGAACATAATTCACCTTTGAGATTCCTATTGAGGAGTTGAA 1293
 Qy 361 TyrIleAlaArgAlaGlyTyrThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
 Db 1294 TATATTGCCAGAGCTGGATGGACTCTGAGGAAAATATGCTTGGTCCATCTCTACTAGAT 1353
 Qy 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
 Db 1354 CGTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATTTATTTATCCCAAGTAGAA 1413
 Qy 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
 Db 1414 GATGATGTTTATGGAAGGACAGAGACTCATTCAGTCAGTCCTGATTCGTGACGCCACTA 1473
 Qy 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
 Db 1474 ATTATCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTT 1533
 Qy 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
 Db 1534 CCCCAGAGTCAGAGAGGAATTTGAGTTATTTTGGCTCTGATGTCGCAAAACAGGTTTC 1593
 Qy 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
 Db 1594 CGTCAITTTATACAAATTTACATCTATTTTAAAGGAACAAATATAAACATCCAGTGGT 1653
 Qy 481 GlyLeuProAlaProSerAspPheLysCysProLysLeuGluIleAlaIleThrSer 500
 Db 1654 GGGCTGCTGCTCCAAAGTATTTCAAGTGTCTATCAAGAGGAGATAGCAATATACCAGT 1713
 Qy 501 GlyGluThrGluValLeuGluValArgHisGlySerAsnIleGlnValAspGluValArgArg 520
 Db 1714 GGTCAATGGGAAGTCTTGGCCGGCATGATCTAATATATCCAAAGTTGATGAGTCAGAGG 1773
 Qy 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer 540
 Db 1774 CTGGTATATTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTAGTAGTCACT 1833
 Qy 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560

1834 TACGTAATCCTGGAGGCTGACAAAGGCTGACTGACCGTGGCTACTCACATCTCTTGCTGC 1893
Qy 561 ILeSerGlnHisCysAspPhePheIleSerLysTyrSerAenGlnLysAenProHisCys 580
Db 1894 ATCAGTCAGCACTGTGACTCTCTTTATAGTAGTATAGTACCAAGAGAGATCCACACTGT 1953
Qy 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
Db 1954 GTGTCCTTTACAAAGCTATCAGTCTCGAATGATGACCACTTTCMAAACAAGAGATTT 2013
Qy 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePhe 620
Db 2014 TGGGCCACCAATTTGGATTGAGT----- 2036
Qy 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
Db 2036 ----- 2036
Qy 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyProGlnValGlnLeu 660
Db 2037 -----CCTCAGGTGCAGTTG 2051
Qy 661 ValAenAenArgPheLysGlyValLysTyrPheArgLeuAenThrLeuAlaSerLeuGly 680
Db 2052 GTGATATATCGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTAGCCTCTCTAGGT 2111
Qy 681 TyrValValValValIleAspAenArgGlySerCysHisArgGlyLeuLysPheGluGly 700
Db 2112 TATGTGTTGTAGTGATAGACACAGGGGATCTGTCCAGGGGCTTAAATTTGAAGGC 2171
Qy 701 AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr 720
Db 2172 GCCTTTAAATATAAATGGTCAATAGAAATTTGACGATCAGGTGGAAGGACTCCAAATAT 2231
Qy 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyr 740
Db 2232 CTAGCTTCTCGATATGATTTTCATTGACTTAGATCGTGTGGCATCCAGCGCTGCTCTAT 2291
Qy 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
Db 2292 GGAGGATACCTCTCCCTGATGGCATTAATGAGAGGTCAGATATCTTCAGGTGCTATT 2351
Qy 761 AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780
Db 2352 GCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACAGGAAGCTTATATG 2411
Qy 781 GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
Db 2412 GGTCACTCTGACCAATGAAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAA 2471
Qy 801 LysPheProSerGluProAenArgLeuLeuLeuLeuHisGlyPheLeuAspGluAenVal 820
Db 2472 AAGTTCCCTCTGNAACCAATCGTTTACTGCTCTTACATGCTTTCTCTGGATGAGAATGC 2531
Qy 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
Db 2532 CAITTTGACATACCACTATATTACTAGATTTTTTAGTGGGGCTGGAAAGCCATATGAT 2591
Qy 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
Db 2592 TTACAGATCTATCTCTCAGGAGACACAGCATAGAGTTCCTGAATCGGAGAACATTAT 2651
Qy 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAenLeuGlySerArgIleAlaAlaLeuLys 880
Db 2652 GAATGTCATCTTTTGCACTACCTTCAAGAAAACCTTGATCAGCTATTGCTGCTCTAAA 2711
Qy 881 ValIle 882
Db 2712 GTGATA 2717
RESULT 13
ABK83331
ID ABK83331 standard; cDNA; 4676 BP.

XX AC ABK83331;
XX 12-AUG-2002 (first entry)
XX cDNA encoding human DPRP-1 splice variant #7.
XX Human; serine protease; dipeptidyl peptidase IV-related protein; DPRP;
KW DPPIV; infection; human immunodeficiency virus; HIV-1; HIV-2; pain;
KW diabetes; infertility; obesity; anorexia; Parkinson's disease; stroke;
KW heart failure; hypertension; urinary retention; osteoporosis; cancer;
KW ulcer; allergy; cancer; psychotic disorder; neurological disorder;
KW dyskinesia; reproductive disorder; inflammatory disorder;
KW metabolic disorder; gene; ss.
XX Homo sapiens.
OS
XX WO200231134-A2.
XX 18-APR-2002.
XX 12-OCT-2001; 2001WO-US031874.
XX 12-OCT-2000; 2000US-0240117P.
XX (FERR) FERRING BV.
XX Qi S, Akinsanya KO, Riviere PJ, Junien J;
XX WPI; 2002-444178/47.
XX P-PSDB; ABG61600.
XX New dipeptidyl peptidase IV-related proteins and nucleic acids encoding
PT the proteins, useful for treating e.g. fungal, bacterial, protozoan and
PT viral infections, cancers, allergies, neurological disorders, or pain.
XX Disclosure; Page 72-73; 113pp; English.
XX The present invention relates to the isolation of novel human serine
CC proteases referred to as dipeptidyl peptidase IV (DPPIV)-related proteins
CC (DPRP). The dipeptidyl peptidase IV-related proteins (DPRP) and nucleic
CC acids encoding them are useful for treating infections such as fungal,
CC bacterial, protozoan and viral infections, particularly infections caused
CC by human immunodeficiency virus (HIV-1 or HIV-2), pain, diabetes,
CC precocious puberty, infertility, obesity, anorexia, bulimia, Parkinson's
CC disease, acute heart failure, hypotension, hypertension, urinary
CC retention, osteoporosis, angina pectoris, stroke, ulcers, asthma,
CC allergies, cancers, migraine, vomiting, psychotic and neurological
CC disorders (e.g. anxiety, dementia, or schizophrenia), and dyskinesias.
CC These may also be used in discovering therapeutic agents for the
CC treatment of reproductive, inflammatory and metabolic disorders. ABK83322
CC -ABK83343 encode human DPRP proteins
XX SQ Sequence 4676 BP; 1424 A; 859 C; 979 G; 1414 T; 0 U; 0 Other;
Alignment Scores: Length: 4676
Pred. No.: 0 Matches: 831
Score: 4385.00 Conservative: 0
Percent Similarity: 94.2% Mismatches: 1
Best Local Similarity: 94.2% Indels: 51
Query Match: 6 Gaps: 1
DB:
US-10-825-632-1 (1-882) x ABK83331 (1-4676)
Qy 1 MetaAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
Db 214 ATGGCAGCAGCAATGGAAACAGACAGCTGGGTGTGAGATATTGAACTCGGACTGT 273
Qy 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
Db 274 GAGGAGAAATATTGAATCACAGGATCGCCCTAAATTTGGAGCCCTTTTATGTTGAGCGGTAT 333

QY 41 SerTyrSerGlnLeuLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
DB 334 TCCTGGAGTCAGCTTAAAGAGCTGCTTGGCCGATACAGAAAATATCATGGCTACATGATG 393
QY 61 AlaLysAlaProHisAspMetPheValLysArgAsnAspProAspGlyProHisSer 80
DB 394 GCTAAGGCCACCATGATTTTCATGTTGTGAGAGGAGATGATCCAGATGGACCTCATTC 453
QY 81 AspArgLysLeuTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
DB 454 GACAGAACTTATTACCTTGGCCATGCTGCTGGTGAGAACAGAGAAAATACACTGTTTATTCT 513
QY 101 GluLysProLysThrIleAsnArgAlaValLeuMetLeuSerTrpLysProLeuLeu 120
DB 514 GAAATTCCTCAAAATATCAATAGAGCAGCAGTCTTAATGCTCTCTTGGAGGCTCTTTTG 573
QY 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140
DB 574 GATCTTTTTCAGGCCAACACTGGACTATGGATATGATTTCTCGAGAGAGAACTATTAGA 633
QY 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
DB 634 GAAAGAAACGCAATTTGGAACAGTCGGAATTCCTTTACGATTTATCACCAGAGAGTGA 693
QY 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
DB 694 ACATTTCTGTTTCAAGCGGTAGTGAATTTATCACGTAAAGATGGAGGCCACAGGA 753
QY 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
DB 754 TTTAGCGAACACCTTTTAAGGCCCAATCTAGTGAACACTAGTTGTCCCAACATACGGATG 813
QY 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
DB 814 GATCCAAATATGTCCTGCTGATCCAGACTGGATGCTTTTATACATAGCAACGATATT 873
QY 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu 240
DB 874 TGGATATCTAACATCGTAACAGAGAGAGAGAGAGACTCATTTATGTGCACATGAGCTA 933
QY 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu 260
DB 934 GCCAACATGGAGAGAGATGCCAGATCAGCTGAGTGCCTACCTTTGTTCTCCAGAGAGAA 993
QY 261 PheAspArgTyrSerGlyTyrTrpCysProLysAlaGluThrThrProSerGlyGly 280
DB 994 TTTGATAGATATCTCGGCTATTTGGTGTGCCAAAGCTGAACACTCCAGTGGTGT 1053
QY 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHisVal 300
DB 1054 AAAATCTTAGAATTTCTATATGAGAGAAATGATGAATCTGAGGTGGAAATTTATCATGTT 1113
QY 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320
DB 1114 ACATCCCTTATGTTGGAACAAAGAGGGCAGATTCATTCCGTTTATCTCAAAAACAGGTACA 1173
QY 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
DB 1174 GCAATCTCTTAAAGTCACATTTTAAGATGTCAGAAATAATGATGATGCTGAAGAGGATC 1233
QY 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
DB 1234 ATAGATGTCATAGATAAGAACTAATTCACCTTTTGAGATTTCTATTGAAGGAGTTGAA 1293
QY 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
DB 1294 TATATTGCCAGAGCTGGATGACTCCTGAGGAGAAATATGCTGGTCCATCTCTACCTAGAT 1353
QY 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
DB 1354 CGCTCCACAGCTCGCTACAGATAGTGTGATCTCACCTGAATTTATTTATCCAGTAGAA 1413
QY 401 AspAspValMetGluArgGlnArgMetGluLeuIleGluSerValProAspSerValThrProLeu 420

DB 1414 GATGATGTTATGGAAGGACAGAGACTCATTCAGTCAGTGCCTGATTCTGTGACGCCACTA 1473
QY 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
DB 1474 ATTATCTATGAAGAAACACACACATCTGGATAAATATCCATGACATCTTTCATGTTTTT 1533
QY 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
DB 1534 CCCAAAGTCAAGAGAGAAATGAGTTTATTTTGGCTCTGAATGCAAAACAGGTTTC 1593
QY 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
DB 1594 CGTCATTTATACAAATTTACATCTATTTTAAGGAAAGCAAAATATAAACGATCCAGTGT 1653
QY 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 500
DB 1654 GGGCTCGCTGCTCCAAGTGATTTCAAGTGTCTATCAAGAGGAGATAGCAATTTACCAGT 1713
QY 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArg 520
DB 1714 GGTGAATGGGAAGTTCTTGGCCGCGATGATCTAATATCCAAAGTTGATGAAGTCAGAAG 1773
QY 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer 540
DB 1774 CTGGTATATTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACTGTACGTAGTCAGT 1833
QY 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
DB 1834 TACGTAATCTCGAGAGGCTGACAAAGGCTGACTGACCGTGGCTACTCACATTTCTTGTGC 1893
QY 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
DB 1894 ATCAGTCAGCACTGTGACTTCTTTATAAGTAGATATAGTACCAGAGAAATCCACACTGT 1953
QY 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
DB 1954 GTGTCCCTTTTACAAGCTATCAAGTCTGAAGATGACCACTTGCAGAAACAAAGAAATTT 2013
QY 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 620
DB 2014 TGGGCCACCATTTTGGATTCAGCAGTCTCTCTCTGACTATACTCTCCAGAAATTTTC 2073
QY 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
DB 2074 TCTTTTGAAGTACTACTGATTTTACATTTGATGGATGCTCTCAAGCCCTCATSATCTA 2133
QY 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 660
DB 2134 CAGCCTGGAAAGAAATATCTCTACTGTCTGTTCATATATGTTGGTGTG----- 2180
QY 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680
DB 2180 ----- 2180
QY 681 TyrValValValValIleAspAsnArgLysSerCysHisArgGlyLeuLysPheGluGly 700
DB 2180 ----- 2180
QY 701 AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr 720
DB 2181 -----GGTCAAAATAGAAAATTTGCGATCAGGTGGAGGAGCTCCAATAT 2222
QY 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyr 740
DB 2223 CTAGCTTCTCGATATGATTTTCAATGACTGTAGATGCTGGGCATCCACGCTGTCTCTAT 2282
QY 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
DB 2283 GGAGGATACCTCTCCCTGATGTCATTAATGCAGAGGTGAGATATCTTTCAGGGTGTCTATT 2342
QY 761 AlaGlyAlaProValThrLeuThrIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780

Db 2343 GCTGGGGCCCGACGTCCTGTGGATCTCTATGATACAGGATACACGGAACGTTATATG 2402
Qy 781 GlyHisProAspGlnAnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
Db 2403 GGTACCCCTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGCCATGCAACAGAA 2462
Qy 801 LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAnVal 820
Db 2463 AAGTTCCCTCTGAACCAATCGTTTACTCTCTTACATGGTTCCTGTGATGAGAATGTC 2522
Qy 821 HisPheAlaHisThrSerIleLeuLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
Db 2523 CATTTTGCATACACAGTATATTACTGAGTGTCTTTTGTAGTGAGGCTGGAAGCCATATGAT 2582
Qy 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
Db 2583 TTACAGATCTATCTCTCAGGAGACACAGACATAGAGTTCCTGAATCGGGAGACATTTAT 2642
Qy 861 GluLeuHisLeuHisTyrLeuGlnGluAnLeuGlySerArgIleAlaAlaLeuLys 880
Db 2643 GAACTGCATCTTTTGCACTACCTTCAAGAAAACCTTGGATCAGTATTGCTGCTAAAA 2702
Qy 881 Vallile 882
Db 2703 GTGATA 2708
RESULT 14
ACA92425
ID ACA92425 standard; DNA; 2929 BP.
XX
AC ACA92425;
XX
DT 15-JUL-2003 (first entry)
XX
DE DNA encoding human PMMM-10.
XX
KW Human; protein modification and maintenance molecule; PMMM; cancer;
KW cell proliferation disorder; atherosclerosis; neurological disorder;
KW epilepsy; Huntington's disease; stroke; immune disorder; allergy;
KW inflammatory disorder; AIDS; developmental disorder; hypothyroidism;
KW Cushing's syndrome; gastrointestinal disorder; epithelial disorder;
KW infection; cytotoxic; antileukosclerotic; anticonvulsant; nootropic;
KW neuroprotective; cerebroprotective; anti-HIV; antiallergic; vulnary;
KW antiinflammatory; thyromimetic; gene; ds.
XX
OS Homo sapiens.
XX
PN WO2003031939-A2.
XX
PD 17-APR-2003.
XX
PF 11-OCT-2002; 2002WO-US032850.
XX
PR 12-OCT-2001; 2001US-0329689P.
PR 25-OCT-2001; 2001US-0335703P.
PR 09-NOV-2001; 2001US-0348887P.
PR 28-NOV-2001; 2001US-0334145P.
PR 06-DEC-2001; 2001US-0337451P.
PR 14-DEC-2001; 2001US-0340584P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Ramkumar J, Gorvad AE, Baughn MR, Emerling BM, Yang J, Lee SY;
PI Tran UK, Becha SD, Duggan BM, Lee EA, Griffin JA, Li JX;
PI Sprague WW, Hafalia AJA, Chawla NK, Lehr-Mason PM, Kable AE, Yue H;
PI Marquis JP, Yao MG, Richardson TW, Tang TY, Jin P, Chien D;
PI Bhatia U, Burrill JD, Lee S, Blake JJ, Ho A, Zheng W;
XX
DR WPI; 2003-430274/40.
DR P-PSDB; ABU92030.
XX
XX New human protein modification and maintenance molecules (PMMM), useful
PT for diagnosing, treating and preventing diseases or conditions associated

PT with the aberrant PMMM expression e.g. cancer, atherosclerosis, or
PT infections.
XX
XX Claim 5; Page 289; 31pp; English.
XX
CC The present invention relates to the isolation of human protein
CC modification and maintenance molecules (PMMM), and the polynucleotide
CC sequences encoding them. A total of 40 PMMM polypeptides (designated PMMM
CC -1 to PMMM-40) are disclosed. The sequences of the invention are useful
CC for diagnosing a condition or disease associated with the expression of
CC PMMM in a subject, preparing a polyclonal or monoclonal antibody, and
CC generating an expression profile of a sample containing the
CC polynucleotides. The diseases or conditions associated with decreased
CC expression or overexpression of PMMM are cell proliferation disorders
CC (e.g. cancer, atherosclerosis), neurological disorders (e.g. epilepsy,
CC Huntington's disease, stroke), immune/inflammatory disorders, (e.g. AIDS,
CC allergies), developmental disorders (e.g. hypothyroidism, Cushing's
CC syndrome), gastrointestinal or epithelial disorders, and infections. The
CC PMMM polypeptides or their fragments are useful in screening compounds
CC for effectiveness as agonists or antagonists of the polypeptides, or in
CC altering the expression of the target polynucleotide and compounds that
CC specifically bind to, or modulate the activity of the polypeptide.
CC ACA92416-ACA92455 encode the human PMMM polypeptides of the invention
XX
SQ Sequence 2929 BP; 864 A; 606 C; 661 G; 798 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 0 Length: 2929
Score: 4353.00 Matches: 824
Percent Similarity: 93.4% Conservative: 0
Best Local Similarity: 93.4% Mismatches: 0
Query Match: 92.6% Indels: 58
DB: 10 Gaps: 1
US-10-825-632-1 (1-882) x ACA92425 (1-2929)
Qy 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
Db 203 ATGGCAGCAGCATGGAACAGACAGCTGGTGTGAGATATTGAATCGGACTGT 262
Qy 21 GluGluAnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
Db 263 GAGGAGAAATTGAAATCAGAGATCGGCTAAATGGAGCCTTTTATGTAGCCGTAT 322
Qy 41 SerTyrSerGlnLeuLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
Db 323 TCTGGAGTCAGCTTAAAGAGCTGCTGCCGATACCAAGAAATATCATGCTACATGATG 382
Qy 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
Db 383 GCTAAGGCACCATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTTCA 442
Qy 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAnArgGluAnThrLeuPheTyrSer 100
Db 443 GACAGATCTATTACCTTGCCTGCTGAGAACAGAGAAATACACTGTTTATTCT 502
Qy 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTyrPlysProLeuLeu 120
Db 503 GAAATTCACCAAACTATCAATAGACAGACAGCTTATGCTCTCTTGAAGCCCTCTTTTG 562
Qy 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140
Db 563 GATCTTTTTCAG----- 574
Qy 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
Db 574 ----- 574
Qy 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
Db 574 ----- 574
Qy 181 PheThrGlnGlnProLeuArgProAnLeuValGluThrSerCysProAsnIleArgMet 200

Db 575 -----CAACACCTTTTAAGGCCCAATCTAGTGAATACTAGTTGTCCCAACATACGGATG 628
QY 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
Db 629 GATCCAAATATATGCTCCCTGCTGATCCAGACTGGATGCTTTTATACATAGCAACGATATT 688
QY 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTrpValHisIleAsnGluLeu 240
Db 689 TGGATATCTAACATCGTAACAGAGAGAAAGAGAGACTCACTTATGTGCACATGAGCTTA 748
QY 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGluGlu 260
Db 749 GCCAACATGGGAAGAGATGCCAGATCAGCTGGAGTCGCTACCTTGTCTCCAAAGAGAA 808
QY 261 PheAspArgTrpSerGlyTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280
Db 809 TTTGATAGATATCTGGCTATTTGGTGGTCCAAAGCTGAACCACTCCAGTGGTGT 868
QY 281 LysIleLeuArgIleLeuTrpGluGluAsnAspGluSerGluValGluIleHisVal 300
Db 869 AAAATCTTTAGAAATCTATATGAAGAAAATGATGAATCTGAGGTGGAAAATTTATTCATGTT 928
QY 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTrpProLysThrGlyThr 320
Db 929 ACATCCCTCATGTTGGAAACCAAGAGGCGCAGATTCATTCCGTTATCTTAAACAGAGTACA 988
QY 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
Db 989 GCNAATCTCTAAAGTCACCTTTTAAAGTGTACAGATGTACAGAAATATGATGATGCTGAGGAAGGATC 1048
QY 341 IleAspValIleAspLysLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
Db 1049 ATAGATGTCATAGATAAGGAACTAATTCACCTTTTTCAGATTCCTTTGAAGGAGTTGAA 1108
QY 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTrpAlaTrpSerIleLeuAsp 380
Db 1109 TATATGTCAGAGCTGGATGGACTCTCTGAGGAAAATATGCTGGTCCATCTCTACTAGAT 1168
QY 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
Db 1169 CGTCCACAGCTCGCTCAGATAGTGTGATCTCACCCTGAATTTATTTATCCAGCTAGAA 1228
QY 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
Db 1229 GATGATGTTATGGAAAGCAGAGACTCATTGAGTCAGTGCCTGATTCCTGAGCGCCACTA 1288
QY 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
Db 1289 ATTATCTATGAAGAAACAACACACATCTGGATAAATATCCATGACATCTTTTCATGTTTT 1348
QY 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
Db 1349 CCCCAAAGTCACGAAGAGAAATGAGTTATTTTTCCTCTGATGCAAAAACAGGTTTC 1408
QY 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTrpLysArgSerSerGly 480
Db 1409 CGTCATTTATACAAATATACATCTATTTTAAAGGAAGCAAAATATAAACCATTCCAGTGGT 1468
QY 481 GlyIleProAlaProSerAspPheLysCysProIleLysGluIleAlaIleThrSer 500
Db 1469 GGCTCCCTGCTCCAAAGTATTTCAAGTGTCTATCAAGAGGAGATAGCAATTTACCAGT 1528
QY 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
Db 1529 GGTGAATGGGAAGTTCTTGGCCGGCATGGATCTAATATCCAAAGTTGATGAAGTCAGAAGG 1588
QY 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer 540
Db 1589 CTGGTATATTTTGAAGGACCAACAGACTCCCTTTTAGGCATACCTGTAGCTAGTACAGT 1648
QY 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560

Db 1649 TAGTAAATCTCTGGAGAGGTGACAAGGCTGACTGACCGTGCTACTCACATTTCTTGCTGC 1708
QY 561 IleSerGluHisCysAspPhePheIleSerLysTrpSerAsnGlnLysAsnProHisCys 580
Db 1709 ATCAGTTCAGCACTGTGACTCTTTTATAAGTAAGTATAGTAACCAAGAGAATCCACACTGT 1768
QY 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
Db 1769 GTGTCCCTTTACAGCTATCAAGTCTCAGATGACCCACTTGCACAAACAAAGGAATTT 1828
QY 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTrpThrProGluIlePhe 620
Db 1829 TGGGCCACCAATTTTGGATTCAGCAGTCTCTCTCTGACTATATCTCTCCAGAAATTTTC 1888
QY 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
Db 1889 TCTTTTGAAGTACTACTGGATTTTACATTTGATGGATGCTCTCAAGCCCTCATGATCTA 1948
QY 641 GlnProGlyLysLysTrpProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 660
Db 1949 CAGCTCGGAAGAAATATCTTACTGTCTGTTTATATATGTTGTTCTCAGTGCAGTTG 2008
QY 661 ValAsnAsnArgPheLysGlyValLysTrpPheArgLeuAsnThrLeuAlaSerLeuGly 680
Db 2009 GTGAATTAATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCTAGCCTCTCTAGT 2068
QY 681 TyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 700
Db 2069 TATGTGTTGTAGTAGACACAAACAGGGGATCTGTCTACCCGAGGGCTTAAATTTTGAAGGC 2128
QY 701 AlaPheLysTrpLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr 720
Db 2129 GCCTTTTAAATATAAATAGGTTCAATAGAAATGACGATCAGGTGGAGGACTCCAATAT 2188
QY 721 LeuAlaSerArgTrpAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyr 740
Db 2189 CTAGCTTCGATATGATGTTTTCATTCAGTTAGATCGTGTGGCATCCACGGCTGCTCAT 2248
QY 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
Db 2249 GGAGGATACCTCTCCCTGATGGCATTAATGCAGAGGTCAAGATATCTTCAGGGTTGCTATT 2308
QY 761 AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780
Db 2309 GCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGAAACGTTATATG 2368
QY 781 GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
Db 2369 GGTCACCTCGACCAAGANTGAACAGGCTATTTACTTAGGATCTGTGGCCATCCAGACGAA 2428
QY 801 LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
Db 2429 AAGTTCCTCTGAAACAAATCGTTACTGCTCTTACATGGTTTCTCGATGAGAATGTC 2488
QY 821 HisPheAlaHisThrSerIleLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
Db 2489 CATTTTGCACATACCAAGTATATTTACTGAGTTTTTTAGTGGGGCTGGAAACCATATGAT 2548
QY 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
Db 2549 TTACAGATCTATCTCCGAGAGACACACGATTAAGGTCTCTGAATCGGGAGACATATAT 2608
QY 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLys 880
Db 2609 GAAGTGCATCTTTTGGCACTACCTTCAAGAAAACCTTGGATCAGCTATTGCTGCTCTAAA 2668
QY 881-Valile 882
Db 2669 GTGATA 2674

RESULT 15
ADV43982
ID ADV43982 standard; cDNA; 2349 BP.

XX ADVA3982;
 XX 10-MAR-2005 (first entry)
 XX Human psychoneuroendocrinimmune expressed sequence tag SEQ ID NO 1610.
 XX
 XX microarray; psychoneuroendocrinimmune; chronic fatigue;
 KW non-insulin dependent diabetes; allergy; immune disorder; inflammation;
 KW cancer; neoplasm; infection; expressed sequence tag; es.
 XX
 XX Homo sapiens.
 XX
 XX WO2004108899-A2.
 XX
 XX 16-DEC-2004.
 XX
 XX 04-JUN-2004; 2004WO-US017686.
 XX
 XX 04-JUN-2003; 2003US-0475915P.
 XX
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 XX Nicholson A, Vernon SD;
 XX
 XX WPI; 2005-031682/03.
 XX
 XX New microarray comprising probes for genes involved in
 PT psychoneuroendocrinimmune (PNI) activity, useful in diagnosing a
 PT condition associated with PNI activity, e.g., inflammatory or infectious
 PT diseases.
 XX
 XX Claim 1; SEQ ID NO 1610; 254pp; English.
 XX
 XX The invention relates to a new microarray which comprises probes for
 CC genes involved in psychoneuroendocrinimmune (PNI) activity. The
 CC microarray is useful in diagnosing a condition associated with PNI
 CC activity, such as CFS, type-2 diabetes, allergic condition, inflammation,
 CC cancer and infection. The present sequence represents a
 CC psychoneuroendocrinimmune gene expressed sequence tag. Note the
 CC specificatio mentions SEQ ID NO of up to 3314 but only sequences up to
 CC SEQ ID NO 1829 are provided.
 XX
 XX SQ Sequence 2349 BP; 723 A; 459 C; 508 G; 659 T; 0 U; 0 Other;
 XX
 XX Alignment Scores:
 XX Pred. No.: 0 Length: 2349
 XX Score: 4118.00 Matches: 782
 XX Percent Similarity: 88.7% Conservative: 0
 XX Best Local Similarity: 88.7% Mismatches: 0
 XX Query Match: 87.6% Indels: 100
 XX DB: 14 Gaps: 1
 XX
 XX US-10-825-632-1 (1-882) x ADVA3982 (1-2349)
 QY 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
 DB 1 ATGGCAGCAGCAATGGAACACAGACAGCTGGGTGTGAGATATTGAACTGGCGACTGT 60
 QY 21 GluGluAenIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
 DB 61 GAGGAGAAATATTGAATACAGAGATCGGCCTAAATTTGGAGCCTTTTATGTGAGCGGTAT 120
 QY 41 SerTTPSerGlnLeuLysLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
 DB 121 TCCTGGAGTCAGCTTAAAAAGCTGCTTCCGATACCAGAAAATATCATGGCTACATGATG 180
 QY 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
 DB 181 GCTAAGGCACCATGATTTTCATGTTTGTGGAAGAGGAATGATCCAGATGACCTCATTTCA 240
 QY 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAenThrLeuPheTyrSer 100
 DB 1321 CCCCAGAGTACAGGAGGAAATTCAGTTTATTTTTCCTGCTGAATCAAAACAGGTTTC 1380

DB 241 GACAGAATCTATTACCTTGCCTGTCGTGGTGAAGAACAGAAAAATACATGTTTTATTCT 300
 QY 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
 DB 301 GAAATTTCCCAAAATCATCAATAGAGCAGCAGTCTTAATGCTCTCTTGGAGGCTCTTTTG 360
 QY 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuAsp 140
 DB 361 GATCTTTTTTCAGGCACACTCGACTATGGAAATGTAATCTCGAGAGAGAACTATTAAAGA 420
 QY 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
 DB 421 GAAAGAAAACGCATTGGACACAGTCGGAATTCCTCTTACGATTTATCACCAGGAATGGA 480
 QY 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
 DB 481 ACAATTTCTGTTTCAAGCCGGTAGTGGAAATTTATCACGTAAAGATGGAGGGCCCAAGGA 540
 QY 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
 DB 541 TTTAGGCAACACCTTTAAGGCCCAATCTAGTGGAAACTAGTTGTCCCAACATACCGATG 600
 QY 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
 DB 601 GATCCAAAATTTATGCCCTGCTGATCCAGACTGGATTGCTTTTATACATAGCAACGATATT 660
 QY 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu 240
 DB 661 TGGATATCTAACATCGTAACACAGAGAGAGAGACTCATTATGTGCAACAATGAGCTA 720
 QY 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu 260
 DB 721 GCCAACATGGAAGAGATGCCAGATCAGCTGGAGTCCGTACCTTTGTTCTCCCAAGAGAA 780
 QY 261 PheAspArgTyrSerGlyTyrTrpCysProLysAlaGluThrThrProSerGlyGly 280
 DB 781 TTTGATAGATATTCCTGGCTATTGGTGTGTCCTCAAAAGCTGAAACAACCTCCAGCTGGTGT 840
 QY 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300
 DB 841 AAAATCTTTAGAAATCTATATGAAGAAATGAAATGAAATCTGAGGTGGAAATTTATTCAGTT 900
 QY 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320
 DB 901 ACATCCCTTATGTTGGAACCAAGGAGGCGAGATTCAATTCGTTTATCTTAAACAGGTACA 960
 QY 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
 DB 961 GCAAAATCCTTAAAGTCACITTTTAAAGATGTGAGAAATAATGATGATGCTGAAGGAAGGATC 1020
 QY 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
 DB 1021 ATAGATGTCATAGATAAGNACTAATTCACCTTTTGAGATTCTATTGGAAGGAGTTGAA 1080
 QY 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTspSerIleLeuLeuAsp 380
 DB 1081 TATATTGCCAGAGCTGGATGGNCTCTGAGGGAANAATATGCTTGGTCCATCTACTAGAT 1140
 QY 381 ArgSerGlnThrArgGluGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
 DB 1141 CGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATTTATTTATCCCAAGTAGAA 1200
 QY 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
 DB 1201 GATGATGTTATGGAAGGAGAGACTCAATGAGTCAGTGCCTGATTCGTGACGCCACTA 1260
 QY 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
 DB 1261 ATTATCTATGAAGAAACACAGACATCTGGATAAATATCCATGATCATCTTTTCATGTTTT 1320
 QY 441 ProGlnSerHisGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
 DB 1321 CCCCAGAGTACAGGAGGAAATTCAGTTTATTTTTCCTGCTGAATCAAAACAGGTTTC 1380

```
Qy 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
Db 1381 CGTCATTATACAAAATTACATCTATTTAAAGGAAGCAAAATATAAACGATCCAGTGGT 1440
Qy 481 GlyLeuProAlaProSerAspPheLysCysProLysGluLysGluLysIleThrSer 500
Db 1441 GGCGTCCTGCTCCTCAAGTGAATTCAGGTGCTCTCAAAAGAGGAGATAGCAATACCAAGT 1500
Qy 501 GlyGluTyrGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
Db 1501 GGTGAATGGGAAGTTCTTGGCGGCATGGATCTAATATCCCAAGTTGATGAAGTCAGAGG 1560
Qy 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer 540
Db 1561 CTGGTATATTTTGAAGGCCACCAAGACTCCCTCTTAGAGCATCACCTGTACGTAGTCAGT 1620
Qy 541 TyrValAsnProGlyGluValThrArgLeuThrArgLeuThrArgGlyTyrSerHisSerCys 560
Db 1621 TAGCTAAATCCTGGAGAGGTGAAGAGGCTGACTGACCGTGGCTACTCACATTTCTTGCTGC 1680
Qy 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
Db 1681 ATCAGTCAGCACTGTGACTTCTTTATAAGTAAGTATAGTAAACCAAGAAATCCACACTGT 1740
Qy 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
Db 1741 GTGTCCCTTTACAGCTATCAAGTCTGAGATGACCACTTCGAAAACAAAGGAATTT 1800
Qy 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluLysPhe 620
Db 1801 TGGGCCACCACTTTTGGATTCAGCAGGTCCTCTTCCTGACTATACCTCCTCAGAAATTTTC 1860
Qy 621 SerPheGluSerThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
Db 1861 TCCTTTGAAAGTACTACTGGATTCATTTGATGGATGCTCTCAAGCCTCAGATCTA 1920
Qy 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 660
Db 1921 CAGCTGGAAGAAATATCTACTGTGCTGTTTCATATATGGTGGTCTCAG----- 1971
Qy 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680
Db 1971 ----- 1971
Qy 681 TyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 700
Db 1971 ----- 1971
Qy 701 AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr 720
Db 1971 ----- 1971
Qy 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyr 740
Db 1971 ----- 1971
Qy 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
Db 1972 -----GTTGCTATT 1980
Qy 761 AlaGlyAlaProValThrLeuTyrPheTyrAspThrGlyTyrThrGluAtgTyrMet 780
Db 1981 GCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACGTATATG 2040
Qy 781 GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
Db 2041 GGTCACTCCCTGACCAGATGACAGGGCTATTACTTAGGATCTGTGGCCATCGAAGCAGAA 2100
Qy 801 LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
Db 2101 AAGTTCCCTCTGAACCAAAATCGTTTACTGCTCTTACATGGTTTCTCGATGAGATGTC 2160
```

```
Qy 821 HisPheAlaHisThrSerIleLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
Db 2161 CATTTTGCATACACAGTATATTTACTGAGTTTTTTAGTGAGGCTGGAAAGCCATATGAT 2220
Qy 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
Db 2221 TTACAGATCTATCCTCAGGAGAGACACAGCATAGAGTTCTGATCGGAGAACATTAT 2280
Qy 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeuLys 880
Db 2281 GAACTGCATCTTTTGGCACTACCTTCAAGAAAAACCTTGGATCAGCTATTGCTGCTTAAAA 2340
Qy 881 ValIle 882
Db 2341 GTGATA 2346
```

Search completed: April 15, 2006, 00:34:34
Job time : 1298 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 15, 2006, 00:34:42 ; Search time 8334 Seconds
(without alignments)
4951.549 Million cell updates/sec

Title: US-10-825-632-1

Perfect score: 4700

Sequence: 1 MAANMETQLGVETPFTADC.....HLHYLQENLGSRIAALKVI 882

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/abes/ABSSWEB spool/US10825632/runat 14042006 092601 4694/app query.fasta_1
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abes08
-USER=US10825632 @CGN 1.1 5315 @runat 14042006 092601 4694 -NCPU=6 -ICPU=3
-NO MAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

```
EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_hic:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gss1:*
10: gb_gss2:*
11: gb_gss3:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4528.5	96.4	3143	4 AK016546	AK016546 Mus muscu
2	4403	93.7	5517	4 AK029788	AK029788 Mus muscu
3	4277	91.0	2649	10 AY411615	AY411615 Homo sapi
4	4138	88.0	2649	10 AY411616	AY411616 Pan trogl
5	4116.5	87.6	2634	10 AY411617	AY411617 Mus muscu
6	3621	77.0	2292	4 CR609512	CR609512 full-leng
7	2833	60.3	3327	4 AK050021	AK050021 Mus muscu

8	2823	60.1	3457	4	AK078301	AK078301 Mus muscu
9	2820	60.0	3376	4	AK050023	AK050023 Mus muscu
10	2139.5	45.5	3243	4	HSN805448	AL834376 Homo sapi
11	1991.5	42.4	3128	4	CR627380	CR627380 Homo sapi
12	1744	37.1	2767	11	DO053109	DO053109 Homo sapi
13	1728	36.8	1265	4	AF176779	AF176779 Homo sapi
14	1720.5	36.6	1292	4	AF175225	AF175225 Homo sapi
15	1693.5	36.0	1042	3	BM557438	BM557438 AGNCOURT
16	1590	33.8	1013	7	CN647232	CN647232 ILLUMIGEN
17	1575	33.5	1027	7	CN646612	CN646612 ILLUMIGEN
18	1555	33.1	1278	4	AF173382	AF173382 Homo sapi
19	1554	33.1	902	8	DN517062	DN517062 1257572 M
20	1542	32.8	957	5	BQ675006	BQ675006 AGNCOURT
21	1539.5	32.8	910	5	BQ675260	BQ675260 AGNCOURT
22	1505.5	32.0	1041	2	BI084090	BI084090 602869453
23	1462	31.1	857	8	DN519771	DN519771 1262448 M
24	1457	31.0	840	8	DN514280	DN514280 1251358 M
25	1430	30.4	968	5	BQ671635	BQ671635 AGNCOURT
26	1427	30.4	910	5	BX372276	BX372276 BX372276
27	1396	29.7	914	5	BX390898	BX390898 BX390898
28	1391	29.6	1984	4	CR612208	CR612208 full-leng
29	1391	29.6	1992	4	CR603620	CR603620 full-leng
30	1370.5	29.2	853	2	BI223892	BI223892 602941035
31	1354	28.8	1082	5	BU239476	BU239476 603322338
32	1349	28.7	875	8	DN068337	DN068337 JGI_CABD4
33	1343	28.6	753	1	AL040398	AL040398 DKEPZp434A
34	1343	28.6	766	6	CD803232	CD803232 UI-M-GV0-
35	1342	28.6	789	2	BG709118	BG709118 602675382
36	1342	28.6	3726	4	AK039652	AK039652 Mus muscu
37	1325	28.2	1090	3	BM910838	BM910838 AGNCOURT
38	1303	27.7	742	7	CN427229	CN427229 170006002
39	1297	27.6	804	7	CN525619	CN525619 UI-M-HN0-
40	1295	27.6	748	6	CD803197	CD803197 UI-M-GV0-
41	1290	27.4	855	8	CK334658	CK334658 JGI_X2T70
42	1286	27.4	746	5	BY751026	BY751026 BY751026
43	1283.5	27.3	738	7	CN427234	CN427234 170006000
44	1281	27.3	779	8	CN783792	CN783792 UI-M-HY0P
45	1277.5	27.2	985	3	BQ068650	BQ068650 AGNCOURT

ALIGNMENTS

RESULT 1
AK016546
LOCUS
DEFINITION
Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clones:4932434F09 product:DIPEPTIDYL PEPTIDASE 8 homolog [Homo sapiens], full insert sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
AK016546
AK016546.1 GI:12855334
HTC; CAP trapper.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
1
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
10349636
REFERENCE
2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
11042159

REFERENCE
3
Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,

QY	220	IleTrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrrValHisAsnGlu	239
Db	1014	ATTTGGATATCAAACTTCGTAACAGGGAAGCGGAGGATCACATACATGATGAG	1073
QY	240	LeuAlaAsnMetGluGluAspAlaAaGSerAlaGlyValAlaThrPheValLeuGlnGlu	259
Db	1074	CTAGCCAAACATGGAAGAGGATCCAGATCAGCTGGGTGGCCACCTTGTCTTCAGAA	1133
QY	260	GluPheAspArgTyrrSerGlyTyrrTrpCysProLysAlaGluThrThrProSerGly	279
Db	1134	GAATTTGACAGATCTCTGGCTACTCTGGTGTGCTCCCAAGCGGAAGAACTCTCTAGTGT	1193
QY	280	GlyLysIleLeuArgIleLeuTyrrGluGluAsnAspGluSerGluValGluIleHis	299
Db	1194	GGTAATAATCTTAGAATCTCTATGAAGAAATGATGATCTGAGGTGGAGATTAATCAT	1253
QY	300	ValThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrrProLysThrGly	319
Db	1254	GTTACGTCCCATGTTGGAAACCAAGGAGGCGAGATTCCTTTCGTTATCCCAAAACAGGC	1313
QY	320	ThrAlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArg	339
Db	1314	ACAGCAAAACCAAGAGTCACTTTCAAGATGTCGAGATTTGTTGATGCTGAGAGGG	1373
QY	340	IleIleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyVal	359
Db	1374	ATTATAGATGTCATAGATAAAGAACTGGTTCAACCTTTCGAGATTCGTTGAGGAGTT	1433
QY	360	GluTyrrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrrAlaTrpSerIleLeuLeu	379
Db	1434	GAATATATTGCCAGAGCGGATGGACTCCAGAGGGGAAACATGCTGCTCCATCTACTA	1493
QY	380	AspArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProVal	399
Db	1494	GACCGTTCCAGACTCACCTGCAGATAGTTCTGATCTCCCTGAGTTATTTCATCCAGTA	1553
QY	400	GluAspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrPro	419
Db	1554	GAAGATGATGCCATGGACAGACAGACTTATAGAGTCAGTTCTGACTCTGGACACCA	1613
QY	420	LeuIleIleTyrrGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisVal	439
Db	1614	CTGATCATCTCAAGAAACACACAGACATCTGGATAAATATCCACGATATTTTCATGTT	1673
QY	440	PheProGlnSerHisGluGluIleGluPheIlePheAlaSerGluCysLysThrGly	459
Db	1674	TTTCCTCAACTCATGAGATGAATTTGATTTATTTTGGCTCTGATGCAAAACAGGT	1733
QY	460	PheArgHisLeuTyrrLysIleThrSerIleLeuLysGluSerLysTyrrLysArgSerSer	479
Db	1734	TTTCGTCACTGTATAAATCATACATCCATTTTAAAGGAGAGCAAAATATAACCGTCCAGT	1793
QY	480	GlyGlyLeuProAlaProSerAspPheLysCysProLysGluGluIleAlaIleThr	499
Db	1794	GGTGGACTACCTGCCCAAGTGATTTCAAGTCTCATCAAGGAGAAATAACAAATTAACC	1853
QY	500	SerGlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArg	519
Db	1854	AGTGGTGAATGGGAAGTACTTGGCCGGCATGATCTAATATCTGGGTTGATGAAGCCAGA	1913
QY	520	ArgLeuValTyrrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrrValVal	539
Db	1914	AAGCTGTGTACTTTGAAGGCACCAAGACTCTCTTTGGAGCATCACCTGTACGTGACC	1973
QY	540	SerTyrrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrrSerHisSerCys	559
Db	1974	AGTTATGCAAAACCTGGAGAGTGGTGGAGCTGACTGACCGGTGGCTACTCACACTCTCTGC	2033
QY	560	CysIleSerGlnHisCysAspPheIleSerLysTyrrSerAsnGlnLysAsnProHis	579
Db	2034	TGCCTCAGCGGATTTGTGACTTCTTCAAGTAAGTACAGCAACAGAGAATCCACAC	2093
QY	580	CysValSerLeuTyrrLysLeuSerSerProGluAspAspProThrCysLysThrLysGlu	599
Db	2094	TGTGTGTCCCTCTCAAACTCTCAAGTCTCTGAGGATGAGCCAGTTCATAAAACAAAGGAA	2153
QY	600	PheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrrThrProGluIle	619
Db	2154	TTTTGGGCGACCATTTTGGATTTCAGAGGTCTCTTCTGACTACACCCCTCCAGAAATTT	2213
QY	620	PheSerPheGluSerThrThrGlyPheThrLeuTyrrGlyMetLeuTyrrLysProHisAsp	639
Db	2214	TTTTCTTTTGAAGTACTTCTGGATTTCACCTGTATGGAATGTTGTATAGCCCTCATGAC	2273
QY	640	LeuGlnProGlyLysLysTyrrProThrValLeuPheIleTyrrGlyGlyProGlnValGln	659
Db	2274	CTACAACTCGAAGAAATATCCCACTGTGTTTATATATATGTTGCTCCCAAGGTCCAG	2333
QY	660	LeuValAsnAsnArgPheLysGlyValLysTyrrPheArgLeuAsnThrLeuAlaSerLeu	679
Db	2334	CTGGTGAACTTCGGTTTAAAGGAGTCAAGTATTTCCGCTCGAACACCCCTGGGCTCCCTG	2393
QY	680	GlyTyrrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGlu	699
Db	2394	GGTTATGTTGTTGTTGATAGACAAACAGGGGATCTGTCCAGGAGACTTAATTTGAA	2453
QY	700	GlyAlaPheLysTyrrLysMetGlyGlnIleLeuAspAspGlnValGluGlyLeuGln	719
Db	2454	GGCGCTTTTAAATATGAGTGGTCAATAGAAATCGATGATCAAGTGGAGGACTCCAG	2513
QY	720	TyrLeuAlaSerArgTyrrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSer	739
Db	2514	TACCTAGCATCTCAGTATGACTTTCATTCAGTGGATGGGTCATCCACGGCTGGTCC	2573
QY	740	TyrGlyGlyTyrrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAla	759
Db	2574	TATGTTGGCTTACTCTCCCTGATGGCATTAATGCAGAGGTCCGATATCTTCCGGGTGGCT	2633
QY	760	IleAlaGlyAlaProValThrLeuTrpIlePheTyrrAspThrGlyTyrrThrGluArgTyrr	779
Db	2634	ATTGCTGGGCGCCAGTCACCTGTGGATCTTCTATGATACAGGATACCGGAGCGCTAT	2693
QY	780	MetGlyHisProAspGlnAsnGluGlyTyrrTyrrLeuGlySerValAlaMetGlnAla	799
Db	2694	ATGGGTACCCCTGACCAAGAAATGACAGGCTACTACTAGGATCTGTGGCCATCGAAGCG	2753
QY	800	GluLysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsn	819
Db	2754	GAGAAGTTCCCTTCAGAACCAACCGGTTACTCTCTTGGCATGGGTTCTTGGATGAGAAT	2813
QY	820	ValHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrr	839
Db	2814	GTTCACTTTGCACACACCATGATATTTGCTGAGTTTTTATGAGGCTGGAAAGCCATAT	2873
QY	840	AspLeuGlnIleTyrrProGlnGluArgHisSerIleArgValProGluSerGlyGluHis	859
Db	2874	GACTTACAGATCTATCTCAGGAGAGGCACAGCATCAGAGTTCTCTGAGTCTGGAGAACAC	2933
QY	860	TyrGluLeuHisLeuLeuHisTyrrLeuGlnGluAsnLeuGlySerArgIleAlaLeu	879
Db	2934	TATGAATGCACTGCTCCACTACCTTCAGGAGAACCTTGGATCCGATCCGTTGCTGCTG	2993
QY	880	LysValIle 882	
Db	2994	AAAGTGATA 3002	
RESULT 2			
AK029788			
LOCUS			
DEFINITION Mus musculus adult male testis cDNA, RIKEN full-length enriched			
library, clone:493056C15 product:DIPEPTIDYL PEPTIDASE 8 homolog			
[Homo sapiens], full insert sequence.			
ACCESSION AK029788			
VERSION AK029788.1 GI:26081520			
KEYWORDS HTC; CAP trapper.			

SOURCE
ORGANISM
Mus musculus (house mouse)

REFERENCE
AUTHORS
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS
Carninci, P. and Hayashizaki, Y.

TITLE
JOURNAL
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)

REFERENCE
AUTHORS
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.

TITLE
JOURNAL
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)

REFERENCE
AUTHORS
Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitasumi, T., Tachiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kaibiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ogasawa, K., Takata, T., Matsudaira, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A., and Hayashizaki, Y.

TITLE
JOURNAL
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)

REFERENCE
AUTHORS
The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.

TITLE
JOURNAL
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)

REFERENCE
AUTHORS
The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

TITLE
JOURNAL
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)

REFERENCE
AUTHORS
Adachi, J., Aizawa, K., Akimura, T., Azakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Inotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y.

TITLE
JOURNAL
Direct Submission

COMMENT
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

FEATURES
source
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/
Location/Qualifiers
1. .5517
/organism="Mus musculus"
/mol_type="mRNA"

misc_feature
/strain="CS7BL/6J"
/db_xref="FANTOM_DB:4930560C15"
/db_xref="taxon:10090"
/clone="4930560C15"
/sex="male"
/tissue_type="testis"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
2221. .4837
/note="DIPPTIDYL PEPTIDASE 8 homolog [Homo sapiens] (SPTP1Q9BM5, evidence: FASTY, 95.7%ID, 100%length, match=2649)
putative"

ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 5517
Score: 4403.00 Matches: 827
Percent Similarity: 95.5% Conservative: 17
Best Local Similarity: 93.6% Mismatches: 17
Query Match: 93.7% Indels: 23
DB: 4 Gaps: 2
US-10-825-632-1 (1-882) x AK029788 (1-5517)
Qy 1 MetAlaAlaAlaMetGluThrGluGluLeuGlyValGluThrPheGluThrAlaAspCys 20
Db 2248 ATGGCAGCAGCAATGGAAACAGACAGCTGGGTGTCGAGATATTTGAACTGCAGAGTGT 2307
Qy 21 GluGlu---AsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArg 39
Db 2308 GAGGAGGGAATGGTGAATCACAGGACGCCCAACTGGAGCCCTTTTATGTCGAACGG 2367
Qy 40 TyrSerTyrSerGlnLeuLysLysLeuAlaAspThrArgLysTyrHisGlyTyrMet 59
Db 2368 TATTCCTGAGTCACTGAAAGAGCTGCTGCTATACCAAGAAATACCAAGCTACATG 2427
Qy 60 MetAlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHis 79
Db 2428 ATGGCTAAGGCGCCACATGACTTTATGTTGTTAAAGAGGAGCGGATCCAGATGGCCCTCAC 2487
Qy 80 SerAspArgLysTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyr 99
Db 2488 TCAGACAGGGTCTATTACCTTGCCATGTCTGGTGAAGAACAGAGAAATACACTATTTTAT 2547
Qy 100 SerGluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTyrLysProLeu 119
Db 2548 TCTGAAATCCCTAAACCATCAACAGACAGCAGTCTTAATGCTTTCTCGAAGCCCTT 2607
Qy 120 LeuAspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeu 139
Db 2608 TTGGATCTTTTCAGCGCGACACTAGACTATGGATGTATTCTCGAGAGGAAGAGTACTC 2667
Qy 140 ArgGluArgLysArgLysGlyThrValGlyIleAlaSerTyrAspTyrHisGlySer 159
Db 2668 AGAAGAAGAAGAGCGCATTTGGACCTGGGGATCGCAGCTTATGATTACCCAGGAAGC 2727
Qy 160 GlyThrPheLeuPheGlnAlaGlySerGlyLysTyrHisValLysAspGlyGlyProGln 179
Db 2728 GGAACATCTCTGTTTCAGCTGGCAGCGCATTTACCATATCAAGATGAGGAGCCACAT 2787
Qy 180 GlyPheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArg 199
Db 2788 GGATTTACACAACAGACTTTGGCGCCCAATTTAGTGAACATAGTTGTCCTCCATATACGA 2847
Qy 200 MetAspProLysLeuCysProAlaAspProAspTyrIleAlaPheIleHisSerAsnAsp 219
Db 2848 ATGGATCCAAATAATTATGCCCTGCTGATCCAGACTGGATAGCTTTTCATTCACGCAATGAT 2907
Qy 220 IleTrrPileSerAsnIleValThrArgGluGluArgArgLeuThrTyrValHisAsnGlu 239
Db 2908 ATTTGGATATCAAACTCTGTAACCGAGGAAGCGGAGGATCATCATACGACATACATGAG 2967

AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A., Todd, M.A., Tanenbaum, D.M., Civeello, D.R., Lu, F., Murphy, B., Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 2649)
AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A., Todd, M.A., Tanenbaum, D.M., Civeello, D.R., Lu, F., Murphy, B., Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.
FEATURES Location/Qualifiers
source 1..2649
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
gene <1..>2649
 /gene="DPP8"
 /locus_tag="HCM4283"
ORIGIN
 Alignment Scores:
 Pred. No.: 0 Length: 2649
 Score: 4277.00 Matches: 815
 Percent Similarity: 92.4% Conservative: 0
 Best Local Similarity: 92.4% Mismatches: 67
 Query Match: 91.0% Indels: 0
 DB: 10 Gaps: 0

US-10-825-632-1 (1-882) x AY411615 (1-2649)

Qy 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluLeuPheGluThrAlaAspCys 20
 Db 1 ATGGCAGCAGCAATGGAAACAGAACAGCTGGGTGTTGAGATATTTCGAACTCGGAGCTGT 60
 Qy 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
 Db 61 GAGGAGATATTGGAATCACAGGATCGGCCTAAATTTGGAGCCTTTTATGTTGAGCGGTAT 120
 Qy 41 SerTrpSerGlnLeuLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
 Db 121 TCTGGAGTCAGCTTAAAAAGCTGCTTGGCGATACAGAAATATCATCGCTACATGATG 180
 Qy 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
 Db 181 GCTAAGGCACCATGATTTTCATGTTGTGAAGAGGATGATCCAGATGGACCTCATTTCA 240
 Qy 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
 Db 241 GACAGAACTATTATCTTGGCATGTCGTGGTGAGAACAGAGAAATACATGTTTATTTCT 300
 Qy 101 GluileProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
 Db 301 GAAATTTCCCAAACTATCAATAGACGACGAGTCTTAATGCTCTCTTTGGAAGCCTCTTTTG 360
 Qy 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140
 Db 361 GATCTTTTTCAGGCAACACTGGACTATGGAATGATTTCTCGAGAGAGAACTATTAAAGA 420
 Qy 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
 Db 421 GAAAGAAAACCGCATTTGGAAACAGTCGGAATTGCTTCTTTACGATTATCACCAAGGAAGTGA 480
 Qy 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
 Db 481 ACATTTCTGTTTCAAGCCGGTAGTGGAAATTTATCACGTAAAGATCGAGGGCCACAAGGA 540

Qy 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
 Db 541 TTTACGCAACCAACCTTTTAAGGCCCAATCTAGTGGAACTAGTTGTCCCAACATACGATG 600
 Qy 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
 Db 601 GATCCAAATATATGCCCTGCTGATCCAGATCGATTGCTTTTATATATACATACACAGATAT 660
 Qy 221 TrpIleSerAsnIleValThrArgGluGluArgArgLeuThrTyrValHisAsnGluLeu 240
 Db 661 TGGATATCTAACATCGTAACAGAGAAAGAGAGACTCATTATGTGCACATAGAGCTA 720
 Qy 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu 260
 Db 721 GCCAACATGGAGAGAGATGCCAGATCAGCTGGAGTCGTACCTTTGTTCTCCAAAGAA 780
 Qy 261 PheAspArgTyrSerGlyTyrTyrTrpCysProLysAlaGluThrThrProSerGlyGly 280
 Db 781 TTTGATAGATATTCTGGCTATTGGTGGTGTCCAAAAGCTGAAACAACTCCAGTGGTGT 840
 Qy 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300
 Db 841 AAAATTTCTAGAAATTTCTATATGAAGAAATGATGAATCTGAGGTGGAATTTATTCATG 900
 Qy 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320
 Db 901 ACATCCCTATGTTGGAAACAGGAGGCGAGATTCATTCGCTTATCTTAAACAGGTACA 960
 Qy 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
 Db 961 GCAAATCTTAAAGTCACATTTTAAAGTGTGAGAAATATGATGATGCTGAGGAAGAGNN 1020
 Qy 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
 Db 1021 NNN 1080
 Qy 361 TyrIleAlaArgAlaGlyTyrThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
 Db 1081 NNN 1140
 Qy 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
 Db 1141 CGCTCCAGACTCGCCTACAGATAGTGTGATCTCCTCAGCTGAATTTATTTATCCAGTAGAA 1200
 Qy 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
 Db 1201 GATGATGTTATGGAAAGGCAGAGACTCATTGAGTCAGTCGCTGATCTGTGACGCCACTA 1260
 Qy 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
 Db 1261 ATTATCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTT 1320
 Qy 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
 Db 1321 CCCAAAGTCACAAAGAGAAATTTGAGTTTATTTTTCCTCTGAAAGCAAAACAGGTTTC 1380
 Qy 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
 Db 1381 CGTCATTTATACAAATTTACATCTATTTTAAAGGAAGCAATATAACATCCATCGTGGT 1440
 Qy 481 GlyLeuProAlaProSerAspPheLysCysProLysGluLeuIleAlaIleThrSer 500
 Db 1441 GGGCTGCTGCTCCCAANN 1500
 Qy 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
 Db 1501 NNN 1560
 Qy 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer 540
 Db 1561 CTGGTATATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACCTGTACGTAGTCAGT 1620

QY 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
 DB 1621 TACGTAATCTCGAGAGGTGCAAGGCTGACTACCGTGGCTACTACATCTTGCTGC 1680
 QY 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
 DB 1681 ATCAGTCAGCAGCTGTGCTCTCTTATTAAGTAGTAGTAACCAAGGAATCCACACTGT 1740
 QY 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
 DB 1741 GTGTCCCTTTTCAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAGAAACAAAGGAATTT 1800
 QY 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 620
 DB 1801 TGGGCCACCATTTTGGATTCAGCAGGCTCTCTCTGACTATATCTCTCCAGAAATTTTC 1860
 QY 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
 DB 1861 TCTTTTGAAGTACTACTGGAATTACATTTGTATGGGATGCTCTACAGGCTCATGATCTA 1920
 QY 641 GlnProGlyLysLeuTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 660
 DB 1921 CAGCTCGAAAGAAATATCTCTACTGTCTGTTCATATATGTTGGTCTCAGGTGCAAGTTG 1980
 QY 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680
 DB 1981 GTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTTAGCCTCTCTAGT 2040
 QY 681 TyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 700
 DB 2041 TATGTGTTGTAGTGATAGACAAACAGGGGATCTGTCCAGGAGGCTTAAATTTGAAGC 2100
 QY 701 AlaPheLysTyrLysMetGlyGlnIleGluIleAspGlnValGluGlyLeuGlnTyr 720
 DB 2101 GCCTTTAAATATAAATGGTCAATAGAAATTCAGCATCAGGTGGAAGGACTCCAAATAT 2160
 QY 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyr 740
 DB 2161 CTAGCTTCTCGATATGATTTCTATTGATCTAGATCGTGTGGCATCCAGCGTGTCTAT 2220
 QY 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
 DB 2221 GAGGATACCTCTCTCTGATGGCATTAATGACAGAGTTCAGATATCTTCAGGTTGCTATT 2280
 QY 761 AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780
 DB 2281 GCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGATACAGGAACGTTATATG 2340
 QY 781 GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
 DB 2341 GGTCACTCCCTGACAGAAATGAACAGGGCTATTACTAGGATCTGTGGCCATGCAAGCANN 2400
 QY 801 LysPheProSerGluProAsnArgLeuLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
 DB 2401 NNNNNNNNNNNNNNCCAAATCGTTACTGTCTCTTACTGCTTCTTCTGGAATGAGAATGC 2460
 QY 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
 DB 2461 CATTTTGCACATACAGATATATTCTGAGTTTTTTAGTGGGGCTGGAAGCCATATGAT 2520
 QY 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
 DB 2521 TTACAGATCTATCTCTCAGGAGGACACAGCATAGAGTTCTGTAATCTGGAGAACATTAT 2580
 QY 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeuLys 880
 DB 2581 GAATGCACTTTTGGCACTACCTTCAAGAAACCTTGGATCAGCTATGCTGCTCTAAAA 2640
 QY 881 ValIle 882
 DB 2641 GTGATA 2646

RESULT 4

AY4111616 2649 bp DNA linear GSS 16-DEC-2003
 LOCUS Pan troglodytes DPP8 gene, VIRTUAL TRANSCRIPT, partial sequence,
 DEFINITION genomic survey sequence.
 ACCSSION AY4111616
 VERSION AY411161.1 GI:39767584
 KEYWORDS GSS.
 SOURCE Pan troglodytes (chimpanzee).
 ORGANISM Pan troglodytes
 Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
 Hominidae; Pan.
 REFERENCE 1 (bases 1 to 2649)
 AUTHORS Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
 Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
 Perrier, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
 Adams, M.D. and Cargill, M.
 TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
 gene trios
 JOURNAL Science 302 (5652), 1960-1963 (2003)
 PUBMED 14671302
 REFERENCE 2 (bases 1 to 2649)
 AUTHORS Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
 Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
 Perrier, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
 Adams, M.D. and Cargill, M.
 TITLE Direct Submission
 JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA
 COMMENT This sequence was made by sequencing genomic exons and ordering
 them based on alignment.
 FEATURES
 source Location/Qualifiers
 1..2649
 /organism="Pan troglodytes"
 /mol_type="genomic DNA"
 /db_xref="taxon:9598"
 gene <1..>2649
 /gene="DPP8"
 /locus_tag="HCW4283"
 ORIGIN
 Alignment Scores:
 Pred. No.: 0 Length: 2649
 Score: 4138.00 Matches: 791
 Percent Similarity: 89.7% Conservative: 0
 Best Local Similarity: 89.7% Mismatches: 91
 Query Match: 88.0% Indels: 0
 DB: 10 Gaps: 0
 US-10-825-632-1 (1-882) x AY4111616 (1-2649)
 QY 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
 DB 1 ATGGCAGCAGCATGGAACACACAGCTGGGTGTGAGATATTGAACTCGGACTGT 60
 QY 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
 DB 61 GAGGAGATATTGAATCACAGGATCGGCTAAATTTGGAGCCTTTTATTTGAGCGGTAT 120
 QY 41 SerTyrSerGlnLeuLysLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
 DB 121 TCCTGGAGTCAGCTTAAAGAGCTGCTTCGCATATACAGAAATATCATGCTACATGATG 180
 QY 61 AlAlaAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
 DB 181 GCTAAGCCACCATGATTTCATGTTGTGAAGAGGAATGATCAGATGGCCCTCATTC 240
 QY 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
 DB 241 GACAGATCTATTACCTTGNNNNNNGGTGGAACAGAAATACACTGTTTATCTCT 300
 QY 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTyrLysProLeuLeu 120


```

Qy 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
Db 2521 TTACAGATCTATCCTCAGGAGACACACAGATTCCTGAATCTGGAGAACATTAT 2580
Qy 861 GluLeuHisLeuHisTyrLeuGlnGluLeuLeuGlySerArgIleAlaAlaLeuLys 880
Db 2581 GAATGTCATCTTTTGCACTACCTTCAGAAACCTTGGATCAGCTATTGCTCTAATA 2640
Qy 881 Valile 882
Db 2641 GTGATA 2646

RESULT 5
LOCUS AY411617 2634 bp DNA linear GSS 16-DEC-2003
DEFINITION Mus musculus DPP8 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY411617
VERSION AY411617.1 GI:39767585
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2634)
AUTHORS Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 2634)
AUTHORS Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source location/Qualifiers
1..2634
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>2634
/genes="DPP8"
/locus_tag="HCM4283"
gene
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 2634
Score: 4116.50 Matches: 782
Percent Similarity: 90.4% Conservative: 16
Best Local Similarity: 88.6% Mismatches: 78
Query Match: 87.6% Indels: 7
DB: 10 Gaps: 2

US-10-825-632-1 (1-882) x AY411617 (1-2634)
Qy 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
Db 1 ATGGCAGCAGCAATGGAACAGACAGATGGGTGCGAGATATTGAACTGCAGAGTGT 60
Qy 21 GluGlu---AsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArg 39
Db 61 GAGAGGAAATCGTGAATCAGAGGACCGACCAACCTGGAGCCCTTTTATGTGCAACGG 120
Qy 40 TyrSerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMet 59

```

```

Db 121 TATCTTGGAGTCAGCTGAAAGAGCTGCTTGCTGATACCAGAAATACCACGGCTACATG 180
Qy 60 MetAlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHis 79
Db 181 ATGGCTAAGGGCCACATGACCTTATGTTGTAAAGAGCAGGATCCAGATGGCCCTCAC 240
Qy 80 SerAspArgIleTyrTrpLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyr 99
Db 241 TCAGACAGGGTCTATTACCTTGCATGCTGTGTGAGAAACAGAGAAATATACATTTTAT 300
Qy 100 SerGluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeu 119
Db 301 TCTGAATTCCTTAAACCATCAACAGCAGCAGCTTTAATGCTTTCTCGAAGCCCTT 360
Qy 120 LeuAspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeu 139
Db 361 TTGGATCTTTTTCAGGGCAGACATAGACTATGGATGTATTCTCGAGAGGAAGACTACTC 420
Qy 140 ArgGluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySer 159
Db 421 AGAGAAAGAAAGCGCATTTGGAACCGTGGGGATCGCAGCTTATGATTACCCACGAGAGC 480
Qy 160 GlyThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGln 179
Db 481 GGAACATTTCTGTTCAAGCTGGCAGCGCATTTTACCATATCAAGATGGAGGACCAT 540
Qy 180 GlyPheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArg 199
Db 541 GGATTTACACACACGCTTTGGGCCCAATTTAGTGGAACTAGTTGTGCCAATATACGA 600
Qy 200 MetAspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAsp 219
Db 601 ATGGATCCAAAATTTATGCCCTGCTGATCCAGACTGGATAGCTTTCAITTCACAGCAATG 660
Qy 220 IleTrpIleSerAsnIleValThrArgGluGluArgArgLeuThrTyrValHisAsnGlu 239
Db 661 ATTTGGATATCAAACTCTGTAACCGAGGAGAGCGGAGATCAATAGTGTCTTCAAGT 720
Qy 240 LeuAlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu 259
Db 721 CTAGCCCAACATGGAAGAGGATCCAGATCAGCTGGTGTGGCCACTTGTCTTCAAGAA 780
Qy 260 GluPheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGly 279
Db 781 GAATTTGACAGATACCTCTGGTGTCTGCTGCCCAAGCGGAAAGAACTCTCTAGTGT 840
Qy 280 GlyLysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHis 299
Db 841 GGTAAATTCCTTAGAATTCCTATGAAGAAATGATGAATCTGAGGTGGAGATTATTTCAT 900
Qy 300 ValThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGly 319
Db 901 GTTACGTCCCCCATGTTGGAACAGAGGCGGAGATTCCTTGTGTTATCCCAAAACAGGC 960
Qy 320 ThrAlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArg 339
Db 961 ACAGCAAAACCAAGGTCACTTTCAAGATGTGGAGATTGTTGTGTGCTGCAGGAGGG 1020
Qy 340 IleIleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyVal 359
Db 1021 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1080
Qy 360 GluTyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeu 379
Db 1081 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1140
Qy 380 AspArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProVal 399
Db 1141 GACCGTTCCAGACTCACCTGCAGATAGTTCTCGATCTCCCTGAGTTATTTCATCCAGTA 1200
Qy 400 GluAspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrPro 419

```

Db	1201	GAAGATGATGCCATGGACACAGACAGACTTTATAGATCAGTCTTCCTGACCTCTGTGTGACCA	1260
Qy	420	LeuIleIleTyrGluGluThrThrAspIleTrrPileAsnIleHieAspIlePheHisVal	439
Db	1261	CTGATCATCTATGAAGAAACACACAGACATCTGGATAAATATCCAGATATTTTTCATGTT	1320
Qy	440	PheProGlnSerHisGluGluGluIleGluPheIlePheAlaserGluCysIleThrGly	459
Db	1321	TTTTCTCAAACTCATGAAGATGAATTTGAGTTTATTTTTCCTCTGGAATGCAAAACAGGT	1380
Qy	460	PheArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSer	479
Db	1381	TTTCGTCTATCTGTATAAATCACATCCATTTTAAAGAGAGCAAAATATATAACGGTCCAGT	1440
Qy	480	GlyGlyLeuProAlaProSerAspPheLysCysProIleLysGluIleAlaIleThr	499
Db	1441	GGTGGACTACCTGCCCAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	1500
Qy	500	SerGlyGluTrrPgluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArg	519
Db	1501	NN	1560
Qy	520	ArgLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisIleLeuTyrValVal	539
Db	1561	AAGCTGGTGATCTTTGAAGGACCAAGAGACTCTCTTTGGAGCATCAGCTGACGACC	1620
Qy	540	SerTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCys	559
Db	1621	AGTTATGCAAAACCTGGGAAGTGGTGAGGCTGACTGACCGTGGCTACTCACACTCTGTC	1680
Qy	560	CysIleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHis	579
Db	1681	TGCCTCAGCCGCAATTTGTGACTTCTTCATTAAGTAAGTACAGCAACAGAGAAATCCAC	1740
Qy	580	CysValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGlu	599
Db	1741	TGTGTCTCCTCTCAAAACTCTCAAGTCTGAGGATGACCCAGTTCATATAAACAAGGAA	1800
Qy	600	PheTrrPAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIle	619
Db	1801	TTTTGGGCCACCAATTTGGATTCAGCAGGCTCTCTTCTGACTACACCCCTCCAGAAAT	1860
Qy	620	PheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAsp	639
Db	1861	TTTTCTTTTGAAGTACTACTGGATTTACACTGTATGGAATGTTGTATAAGCCTCATGAC	1920
Qy	640	LeuGlnProGlyLysIleTyrProThrValLeuPheIleTyrGlyGlyProGlnValGln	659
Db	1921	CTACAACCTGGAAGAAATATACCCACTGTGTTTATATATATATGTTGTTCCCCAGGTGCAG	1980
Qy	660	LeuValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeu	679
Db	1981	CTGTTGAACATCGGTTTAAAGAGTCAAGTATTTCCGCTGGAACACCTGGCTCCCTG	2040
Qy	680	GlyTyrValValValIleAspAsnArgLysSerCysHisArgGlyLeuLysPheGlu	699
Db	2041	GGTTATGTGGTGTGTATAGAACACAGGGGCTCTGTACCCGAGGACTTAATTTGAA	2100
Qy	700	GlyAlaPheLysTyrLysMetGlyGlnIleAspAspGlnValGluGlyLeuGln	719
Db	2101	GGCGCCTTTAAATATAAATGGGTCAATAGAAATCGATATCAAGTGAAGGACTCCAG	2160
Qy	720	TyrLeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrrPser	739
Db	2161	TACCTAGCATCTCAGTATGACTTTCATTGAGTGGATCGAGTGGCATCCACGGCTGGTCC	2220
Qy	740	TyrGlyGlyTyrLysSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAla	759
Db	2221	TATGGTGGCTTACCTCTCCCTGATGGCATTAATGCAAGAGTCCGATATCTTCCGGGTGGCT	2280
Qy	760	IleAlaGlyAlaProValThrLeuTrrPilePheTyrAspThrGlyTyrThrGluArgTyr	779
Db	2281	ATTGCTGGGGCCCCAGTCAACCTGTGGATCTTTCTATGATACAGGATACACGGAGCGCTAT	2340
Qy	780	MetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAla	799
Db	2341	ATGGGTACCCTTGACCAAGATGAACGGGCTACTACTTAGGATCTGTGSCCATGAAGCA	2400
Qy	800	GluLysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsn	819
Db	2401	-----CCAAACCGGTACTCTCTCTTGATGGGTCTTGGATGAGAT	2442
Qy	820	ValHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyr	839
Db	2443	GTTCACTTTGGACACACAGTATATCTGAGTTTTTAGTGAGGCTCGAAGCCATAT	2502
Qy	840	AspLeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHis	859
Db	2503	GACTTACAGATCTATCTCTCAGGAGAGCAAGATCAGATCTCTGAGCTGGAGAACAC	2562
Qy	860	TyrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeu	879
Db	2563	TATGAACTCACCTGCTCCACTTTCAGGAGAACCTTGGATCGGTATCGCTGCTCTG	2622
Qy	880	LysValIle 882	
Db	2623	AAAGTGATA 2631	
RESULT 6			
CR609512 2292 bp mRNA linear HTC 21-JUL-2004			
LOCUS full-length cDNA clone CS0DL005YD02 of B cells (Ramos cell line)			
DEFINITION Cot 25-normalized of Homo sapiens (human).			
ACCESSION CR609512.1 GI:50490319			
VERSION HTC; CDSLT cDNA			
KEYWORDS Homo sapiens (human)			
SOURCE Homo sapiens			
ORGANISM Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;			
Hominidae; Homo.			
REFERENCE 1 (bases 1 to 2292)			
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.			
TITLE Full-length cDNA libraries and normalization			
JOURNAL Unpublished			
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue 2 (bases 1 to 2292)			
REFERENCE Genoscope.			
AUTHORS Direct Submission			
TITLE Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :			
JOURNAL BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)			
COMMENT - Web : www.genoscope.cns.fr 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.			
FEATURES Location/Qualifiers			
source			
1..2292			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="CS0DL005YD02"			
/issue_type="B cells (Ramos cell line) Cot 25-normalized"			
/plasmid="pCMVSPORT_6"			
ORIGIN			
Alignment Scores:			
Pred. No. : 0			
Score: 3621.00			
Length: 2292			
Percent Similarity: 87.0%			
Matches: 692			
Best Local Similarity: 86.8%			
Conservative: 1			
Query Match: 77.0%			
Mismatches: 0			
Indels: 104			
Gaps: 2			
DB:			

QY 753 SerAspIlePheArgValAlaIleAlaGlyAlaProValThrLeuTrpIlePheTyRasp 772
Db 2402 CCACAGGTGGTTCAGGAGGAGGATTCGGGGGCGCTTCCTGTCACGTGTGGATGCCCTATGAC 2461
QY 773 ThrGlyTyThrGluAArgTyMetGlyHisProAspGlnAsnGluGlnGlyTyTyRleu 792
Db 2462 ACAGGGTACACGGAACACATACATGATGTCCTCCGAAATAAACCACAGGCTATGAGGCA 2521
QY 793 GlySerValAlaMetGlnAlaGluLysPheProSerGluProAsnArgLeuLeuLeu 812
Db 2522 GGGTCTGAGCGCTTCGATGTGGAGAGCTCCCAATGAGCCTAAACGCGCTGCTATCCTC 2581
QY 813 HisGlyPheLeuAspGluAenValHisPheAlaHisThrSerIleLeuLeuSerPheLeu 832
Db 2582 CACGGCTTCCTGGACGAGACGTCATCTTCTCCACACAAATTTCTGCTGTCACGCTG 2641
QY 833 ValArgAlaGlyLysProTyArgPheLeuGlnIleTyProGlnGluArgHisSerIleArg 852
Db 2642 ATCCGAGCAGGAAAGCCATACACGCTTCAGATCTACCCAAACGAGACATAGCATCGC 2701
QY 853 ValProGluSerGlyGluHisTyGluLeuHisLeuLeuHisTyRleuGlnGluAsnLeu 872
Db 2702 TCCCGGAGTCCGGAGAGCATACGAGGTGACGCTGCTGCACCTTCTGCAGGACACCTG 2761
RESULT 8
AK078301
LOCUS 3457 bp mRNA linear HTC 03-APR-2004
DEFINITION Mus musculus adult male olfactory brain cDNA, RIKEN full-length
enriched library, clone:6430584G11 product:DIPEPTIDYL PEPTIDASE 9
homolog [Homo sapiens], full insert sequence.
AK078301 GI:26347124
VERSION HTC; CAP trapper.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Euteheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
1
REFERENCE Carninci, P. and Hayashizaki, Y.
AUTHORS Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
PUBMED 10349636
REFERENCE Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
AUTHORS Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
PUBMED 11042159
REFERENCE Shibata, K., Itoh, M., Aizawa, K., Nagaoaka, S., Sasaki, N., Carninci, P.,
AUTHORS Konno, H., Akiyama, J., Nishi, K., Kitagawa, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Otsu, K., Tanaka, T., Matsumura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
PUBMED 11076861
REFERENCE The RIKEN Genome Exploration Research Group Phase II Team and the
AUTHORS FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
PUBMED 11576861
REFERENCE The FANTOM Consortium and the RIKEN Genome Exploration Research
AUTHORS Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)

6 (bases 1 to 3457)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Haeishizume, W.,
Hayashida, K., Hayashizaki, T., Hiramoto, K., Hiraoka, T., Hirokawa, T.,
Hori, F., Imocani, K., Ishii, Y., Itoh, M., Kagawa, T., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, B., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sugade, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toyota, F., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@sc.riken.jp,
URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.jp/
URL:http://fantom.gsc.riken.jp/
FEATURES
Location/Qualifiers
1..3457
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="GeneID:3457"
/db_xref="FANTOM_DB:6430584G11"
/db_xref="taxon:10090"
/clone="6430584G11"
/sex="male"
/tissue_type="olfactory brain"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
182..2770
/note="unnamed protein product; DIPEPTIDYL PEPTIDASE 9
homolog [Homo sapiens] (SPTRI:AAI47179, evidence: FASTY,
92%ID, 99.6%length, match=2580)
putative"
/codon_start=1
/protein_id="BAC37211.1"
/db_xref="GI:26347125"
/translation="MCSGVSEVQVAAGDMDTAARFCVQKHWGDLRSIHGRSKS
GLVSKPHDFQVKPDESGPHSHRLYLGMYPGRENLSLLYSEIPKVKRKEALLLL
SWKMLDHFQATPHHGVYREELRLERKLGVPGITSYDFHSESLFLPQANSLFLH
CRDGGNGFMVPMKPLEIKTCGSGPRMDPKICPADPAFFSFINNSDLWVANETGEE
RLTFCGSGAGVLNPKSAGVATVIOEEFDRFTGWCPTASWEGSLKTLRLY
EVDSEVEVHVPSPALERKTDYRPTGSKPKIALKLAELQTHQKIVSCE
KELVPFSLFPKVEYIARAGWTRDKYAMFLYRPOQLQVLLFPALFIPAVESE
ACQAAARAVPKNQPFVITYEETVNVINVDIFHPFPAEGQQDFCFRANECKTFQ
CHLYVTVELTKDYMTPELSPTEDEFKCPKEKEVALTSGEWELSRSGKIVWNEQ
TKLVYFQTKDTPLEHLVYVYESAGEI VRLTTLGFSHSCMSQSFDMFVSHYSVS
TPCPVHVYKLSGDDPLHKOPRFWASMEANCPDYVPEIFHFHTRADVQLGMI
YKPHTLQPKRHPTVLFVYGGPQVQVNNSGFKIGIKRLNLASLGVAVVVDGRGSC
ORGLHEGALKNQMGVEIDVEGLOVAKYGFIDLSRAVHSGVSGFLSLMELI
HKPOYKVAIAGAPVTVMAYDTGYTRYMDPENNKOGYAGSVLHVKLEMLNPR
LILALGELDENVHFHTNLQLIRAGKPYQLIYPNERHSIRCRESGHEVETLLH
FLOEHL"

Alignment Scores:	1,02e-311	Length:	3457
Pred. No.:	2823.00	Matches:	510
Score:	76.7%	Conservative:	134
Percent Similarity:	60.7%	Mismatches:	194
Best Local Similarity:	60.1%	Indels:	2
Query Match:			

DB: 4 Gaps: 2
US-10-825-632-1 (1-882) x AK078301 (1-3457)
QY 35 PheTyrValGluArgTyrSerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLys 54
DB 248 TTCTGTGCGAAGACACTCTGGGGATGGGCTGGTAGCATATATCCACGGCAGTGCAG 307
QY 55 TyrHisLysTyrMetMetAlaLysAlaProHisAspPheMetPheValLysArgAsnAsp 74
DB 308 TCCTGGGCGCTCATTTGCAGCAAGGCCCCCGCCAGCACTTCAGTTTGCAGAAAGCTGAC 367
QY 75 ProAspGlyProHisSerAspArgLysTyrTyrLeuAlaMetSerGlyGluAsnArgGlu 94
DB 368 GAGTCTGGCGCCCACTCTCCAGGTCTCTATTACCTCGAATGCTTTACGGCGCGGTGAG 427
QY 95 AsnThrLeuPheTyrSerGluLeuPheProLysThrIleAsnArgAlaAlaValLeuMetLeu 114
DB 428 AACTCCCTCTCTACTCTCGAGATCCCAAGAAAGTGGGAAGGAGGCCCTGCTGCTGTG 487
QY 115 SerTrpLysProLeuLeuAspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArg 134
DB 488 TCCTGAGACGAGTCTGGACCACTTCAGGCCACACCCACCACCATGGTCTACTCCCGA 547
QY 135 GluGluGluLeuLeuArgGluArgLysArgLysIleGlyThrValGlyIleAlaSerTyrAsp 154
DB 548 GAGGAGGAGTACTCGGGAGCGCAAGCGCTGGGGCTCTTCGGAATCACTCTTATGAC 607
QY 155 TyrHisGlnGlySerGlyThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLys 174
DB 608 TTCCACAGTGAAGCGGCTCTCTCTCCAGGCCAGCAATAGCCTGTTCACCTGACGG 667
QY 175 AspGlyGlyProGlnGlyPheThrGlnProLeuArgProAsnLeuValGluThrSer 194
DB 668 GATGTGGCAGAAATGGCTTTATGGTGTCCCGATGAGCCACTGGAGATCAAGACTAG 727
QY 195 CysProAsnIleArgMetAspProLysLeuCysProAlaAspProAspTrpIleAlaPhe 214
DB 728 TGTTCTGGGCCACGATGAGCCCAAAATCTGCCCGCAGACCTGCTCTTTTCTTTC 787
QY 215 IleHisSerAsnAspIleTrpIleSerAsnIleValThrArgGluArgLysLeuThr 234
DB 788 ATCAACAACAGTGATCTGTGGTGGCAACATCGAGACTGGGGAGGAACGGCGGTCAAC 847
QY 235 TyrValHisAsnGluLeuAlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThr 254
DB 848 TTCTGTACAGGGTTACGTGCTCTGGACAAATCCCAATCACAGCGGTGGCCACC 907
QY 255 PheValLeuGlnGluPheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGlu 274
DB 908 TTGTTCATCCAGGAGGATTCGACCGCTTCACTGGGTGCTGGTGGTGGCCCGGCTCT 967
QY 275 ThrThrProSerGlyGly---LysIleLeuArgIleLeuTyrGluGluAsnAspGluSer 293
DB 968 TGGGAAGGCTCCGAAGTCTCAAGAGCTCGCATCTCTATATGAGGAAGTGGACGAGTCT 1027
QY 294 GluValGluIleIleHisValThrSerProMetLeuGluThrArgAlaAspSerPhe 313
DB 1028 GAAGTGGAGGTCATTATGTGCTCCCTCCCGCCCTGGAGGAGGAGGAGCGGACTCTTAC 1087
QY 314 ArgTyrProLysThrGlyThrAlaAsnProLysValThrPheLysMetSerGluIleMet 333
DB 1088 CGCTACCCAGACAGGAGGAGCAAGAACCCCAAGATTGCCCTGAAGTGGGTGAGTCCAG.1147
QY 334 IleAspAlaGluGlyArgIleAlaAspValIleAspLysGluLeuIleGlnProPheGlu 353
DB 1148 ACGGACCATCAGGCAAAATCGTGTCAAGCTGCGAGAAGAACTGTGTACAGCCATTCAGC 1207
QY 354 IleLeuPheGluGlyValGluTyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyr 373
DB 1208 TCCCTTTTCCCAAGTGGAGTACATCGCCGGCTGGCTGGACACGGGAGCAAAATAT 1267
QY 374 AlaTrpSerIleLeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSerPro 393
DB 1268 GCTGGGCATGTTCTTACCGTCCCGCAGACAGCGTTGAGTTCCTTCTTACCCCT 1327
QY 394 GluLeuPheIleProValGluAspValMetGluArgGlnArgLeuIleGluSerVal 413
DB 1328 GCTCTCTTCACTCCCGCGTTCAGAGTGGAGCCCGCAGCGGAGCGAGCTGCCAGAGCGTC 1387
QY 414 ProAspSerValThrProLeuIleIleTyrGluGluThrThrAspIleTrpIleAsnIle 433
DB 1388 CCCAAGATGTGCAGCCCTTTGTCATCTATGAAGAAGTCAACATGTGTGATCAACGCT 1447
QY 434 HisAspIlePheHisValPheProGlnSerHis---GluGluGluIleGluPheIlePhe 452
DB 1448 CACGACATCTCCACCGCTTCTCAGGCTGAGGCCGACAGGAGCTTTGTTCCTTCCT 1507
QY 453 AlaSerGluCysLysThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLysGlu 472
DB 1508 GCCAAGCATCAAGACTGGCTTCTGCCACCTGTACAGGGTGCACAGTGAACCTTAAACC 1567
QY 473 SerLysTyrLysArgSerSerGlyLeuProAlaProSerAspPheLysCysProIle 492
DB 1568 AAGGACTATGACTGGACGGAACCCCTCAGCCCTACAGAAGATGAGTTTAAAGTGCCTC 1627
QY 493 LysGluGluIleAlaIleThrSerGlyLeuTrpGluValLeuGlyArgHisGlySerAsn 512
DB 1628 AAGGAGGAGTGGCTTCCAGCAGTGGGAGTGGAGTCTTGTGCGAGCATGGCTCCAG 1687
QY 513 IleGlnValAspGluValArgArgLeuValTyrPheGluGlyThrLysAspSerProLeu 532
DB 1688 ATCTGGGTCAACGACGACGACGAGCTGGTGTACTTCAAGGTACAAAGGACACACCGCTG 1747
QY 533 GluHisLeuTyrValValSerTyrValAsnProGlyGluValThrArgLeuThrAsp 552
DB 1748 GAACATCACTCTATGTGTGTCAGTCTACGATCAGCAGCGGAGATCGTGCAGCTCAC 1807
QY 553 ArgGlyTyrSerHisSerCysLysSerGlnHisCysAspPhePheIleSerLysTyr 572
DB 1808 CTGGGTCTCCACAGCTGCTCCATGAGCCAGAGCTTCGACATGTTGCTGAGTCACTAC 1867
QY 573 SerAsnGlnLysAsnProHisCysValSerLeuTyrLysLeuSerSerProGluAspAsp 592
DB 1868 AGCAGTGTGAGCAGCGCCCTCTGTGTACATGTGTACAAGCTGAGCGGCCCGATGATGAC 1927
QY 593 ProThrCysLysThrLysGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuPro 612
DB 1928 CCACTGCAACAACAACCCAGCTTCTGGGCCAGCATGTATGAGGAGGAGCAATTTGCC 1987
QY 613 AspTyrThrProProGluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGly 632
DB 1988 GACTATGTGCCCCCTGAGATCTTCCACTTCCACACCCCGTGCAGACGTCTACGGC 2047
QY 633 MetLeuTyrLysProHisAspLeuGlnProGlyLysLysTyrProThrValLeuPheIle 652
DB 2048 ATGATCTACAGCCACACACCTCGAACCTGGAGGAAGACACCCACCTGTGCTTTGTC 2107
QY 653 TyrGlyGlyProGlnValGlnLeuValAsnAsnArgPheLysGlyValLysTyrPheArg 672
DB 2108 TATGGGGGCCACAGGTGTCAGTTGTGTGAACAACCTCTTTAAGGCGATCAATATCTCG 2167
QY 673 LeuAsnThrLeuAlaSerLeuGlyTyrValValValValIleAspAsnArgGlySerCys 692
DB 2168 CTAAATACACTGGCATCTCTGGGCTATGCTGTGGTGGTGTATGATGATGTCGGGGCTCT 2227
QY 693 HisArgGlyLeuLysPheGluGlyValaPheLysTyrLysMetGlyGlnIleGluIleAsp 712
DB 2228 CACGGGGCTGCACTTCGAGGGGGCCCTGAAATAATCAATATGGCCAGGTCGAGATTGAG 2287
QY 713 AspGlnValGluGlyLeuGlnTyrLeuAlaSerArgTyrAspPheIleAspLeuAspArg 732
DB 2288 GACCAAGTGGAGGCTTGGCAGTACCTGGCTGAGAAGTATGCTTCATTGCTTGCAGCGCA 2347
QY 733 ValGlyIleHisGlyTrpSerTyrGlyTyrLeuSerLeuMetAlaLeuMetGlnArg 752

```

Db      2348  GTCGCCATCCATGGCTGGCTCTACGGGGGCTTCTCTCACTCATGGGGCTCATCCACAAG 2407
Qy      753  SerAspIlePheArgValAlaIleAlaGlyAlaProValThrLeuTyrIlePheTyrAsp 772
Db      2408  CCACAAGTGTTCRAAGGTAGCATTCGGGGGCTCTCTGTCTCACTGTGTGGATGGCTATGAC 2467
Qy      773  ThrGlyTyrThrGluArgTyrMetGlyHisProAspGlnAenGluGlnGlyTyrTyrLeu 792
Db      2468  ACAGGGTACACGAGACCATCATCGATGTGTCGCCGAAATACCAACGAGGCTATGAGCA 2527
Qy      793  GlySerValAlaMetGlnAlaGluLysPheProSerGluProAsnArgLeuLeuLeu 812
Db      2528  GGGTCTGTAGCTGTCGATGTGGAGAGCTGCCCAATGACCTTAACCGCTCTGTATCTC 2587
Qy      813  HisGlyPheLeuAepGluAenValHisPheAlaHisThrSerIleLeuLeuSerPheLeu 832
Db      2598  CACGGCTTCCTGGACGAGACGTTCACTTCTTCCACACAAATTTCTGTGTGCCAGCTG 2647
Qy      833  ValArgAlaGlyLysProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleArg 852
Db      2648  ATCCGAGCAGGAAGCCATACACAGCTTCAGATCTACCCAAACGAGAGACATAGCATCCGC 2707
Qy      853  ValProGluSerGlyGluHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeu 872
Db      2708  TCCCGCAGTCCCGAGAGCATTCAGAGTGCAGCTCTCTCACTTCTTCGAGGAACACCTG 2767

RESULT 9
LOCUS   AK050023
DEFINITION Mus musculus adult male liver tumor cDNA, RIKEN full-length
          enriched library, clone:C730003H19 product:DIPEPTIDYL PEPTIDASE 9
          homolog (Homo sapiens), full insert sequence.
ACCESSION AK050023.1 GI:26093756
KEYWORDS HTc; CAP trapper.
SOURCE   Mus musculus (house mouse)
ORGANISM Mus musculus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
          Sciurognathi; Muridae; Murinae; Mus.
1
REFERENCE
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
          Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE    High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
PUBMED  10349636
AUTHORS
TITLE    Normalization and subtraction of cap-trapper-selected cDNAs to
          prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
PUBMED  11042159
AUTHORS
REFERENCE
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
          Konno,H., Akiyama,J., Nishi,K., Kiteunai,T., Tashiro,H., Itoh,M.,
          Sumi,N., Tshii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
          Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kaishiwagi,K.,
          Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
          Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J.,
          Okazaki,Y., Muramatsu,M., Inoue,I., Kira,A. and Hayashizaki,Y.
TITLE    RIKEN integrated sequence analysis (RISA) system--384-format
          sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
PUBMED  11076861
AUTHORS
REFERENCE
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the
          FANTOM Consortium.
TITLE    Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
PUBMED  11500000
AUTHORS
REFERENCE
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research
          Group Phase I & II Team.
TITLE    Analysis of the mouse transcriptome based on functional annotation

```

```

of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 3376)
ADACHI,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
Hayashida,K., Hayatsu,N., Hanagaki,T., Hara,A., Hashizume,W.,
Fukuda,S., Furuno,M., Hiramoto,K., Hiramoto,K., Hiraoka,T., Hirozane,T.,
Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
Kato,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuura,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohnato,N.,
Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
Muramatsu,M. and Hayashizaki,Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Teurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@sc.riken.jp,
URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Tissue was provided by William A. Held, Roswell Park Cancer
Institute, Department of Molecular and Cellular Biology, Elm and
Carlton Streets, Buffalo, NY 14263, whose assistance we gratefully
acknowledge.
please visit our web site for further details.
URL:http://genome.gsc.riken.jp/
URL:http://fantom.gsc.riken.jp/.

FEATURES             Location/Qualifiers
     source            1..3376
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="FANTOM_DB:C730003H19"
                     /db_xref="taxon:10090"
                     /clone="C730003H19"
                     /sex="male"
                     /cell_type="tumor"
                     /tissue_type="liver"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="adult"
     misc_feature      235..2824
                     /note="DIPEPTIDYL PEPTIDASE 9 homolog [Homo sapiens]"
                     (SPTR|AAL47179, evidence: FASTY, 92%ID, 99.6%length,
                     match=2580)
                     putative"

ORIGIN
Alignment Scores:
  Pred. No.:      2,17e-311      Length:      3376
  Score:          2820.00      Matches:    511
  Percent Similarity: 76.7%      Conservative: 134
  Best Local Similarity: 60.8%      Mismatches: 193
  Query Match:    60.0%      Indels:      3
  DB:              Gaps:        2

US-10-825-632-1 (1-882) x AK050023 (1-3376)
Qy      35  PheTyrValGluArgTyrSerTyrSerGlnLeuLysLysLeuLeuAlaAspThrArgLys 54
Db      301 TTCTGTGTGCAGAGCACTGTGGATGGCTGCTAGCATTTATCCACGCGATCCGAC 360
Qy      55  TyrHisGlyTyrMetMetAlaLysAlaProHisAspPheMetPheValLysArgAsnAsp 74
Db      361 TCCTCGGGCCCTCATTTGTCAGCAGGCCGCCACCACTTCCAGTTTGTCAGAGACCTGAC 420

```

QY 75 ProAspGlyProHisSerAspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGlu 94
Db 421 GAGTCTGGCCGCCCACTCTCAACCGTCTATTACTCGGAATGCCCTTACGGCAGCGGTGAG 480
QY 95 AsnThrLeuPheTyrSerGluIleProLysThrIleAsnArgAlaAlaValLeuMetLeu 114
Db 481 AACTCCCTCTCTACTCCGAGATCCCGAAGAGTCCGGAAGAGGCCCTCTGCTGCTG 540
QY 115 SerTrpLysProLeuLeuAspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArg 134
Db 541 TCCTGGAGAGAGTCTGGACCACTTCCAGGCCACACCCACCAGTGTCTACTCCCGA 600
QY 135 GluGluGluLeu-LeuArgGluArgLysArgIleGlyThrValGlyIleAlaSerTyrAs 154
Db 601 GAGGAGAGGTACCTGCGGAGCGCAAGCGCTGGCGCTCTTCGGAATCACTCTTATGA 660
QY 154 pTyrHisGlnGlySerGlyThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLy 174
Db 661 CTTCACAGTGTAGAGCGGCTCTCTCTCTTCCAGGCCAGCAATAGCCTGTTCACCTGCAG 720
QY 174 sAspGlyGlyProGlnGlyPheThrGlnProLeuArgProAsnLeuValGluThrSe 194
Db 721 GATGTGTGGCAAGATGGCTTTATGTGTCCCGATGGAAGCACTGGAGATCAAGACTCA 780
QY 194 rCysProAsnIleArgMetAspProLysLeuCysProAlaAspProAspTrpIleAlaPh 214
Db 781 GTGTCTGGGCCACGCATGGACCCCAAAATCTGCCCGCAGACCCCTCTCTTTTCCTT 840
QY 214 eIleHisSerAsnAspIleTrpLysSerAsnIleValThrArgGluGluArgArgLeuTh 234
Db 841 CATCAACAACAGTATCTGTGGTGGCAACATCGAGACTGGGGAGGAACCGCGGCTCAC 900
QY 234 rTyrValHisAsnGluLeuAlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaIaTh 254
Db 901 CTTCCTGTACACAGGGTTCAGTGTCTGTGGCAATCCCAATCAGCAGGCGTGGCCAC 960
QY 254 rPheValLeuGlnGluPheAspArgTyrSerGlyTyrTrpTyrCysProLysAlaG1 274
Db 961 CTTTGTATCCAGGAGGATTCGACCGCTTCACTGGGTGTCTGTGGTGGCCCGGCTC 1020
QY 274 uThrThrProSerGlyGly---LysIleLeuArgIleLeuTyrGluLeuAsnAspGluSe 293
Db 1021 TTGGAGAGGCTCCCAAGGTCTCAAGACGCTGCGCATCTATATGAGGAAGTGGACGATC 1080
QY 293 rGluValGluIleHisValThrSerProMetLeuGluThrArgArgAlaAspSerPh 313
Db 1081 TGAAGTGGAGGTCTATCTATGCTCCCTCCCGGCTGGAGAGAGACCGACTCTCTA 1140
QY 313 eArgTyrProLysThrGlyThrAlaAsnProLysValThrPheLysMetSerGluIleMe 333
Db 1141 CCGCTACCCAGGACAGGACGACGAACCCCAAGATTCCTGMAAGTGGCTGAGCTCCA 1200
QY 333 tIleAspAlaGluGlyArgIleIleAspValIleAspLysGluLeuIleGlnProPheG1 353
Db 1201 GACGGACCATCAGGCGCAAAATCGTCAAGCTCGGAGAAAGAACTGGTACAGCCATTGAG 1260
QY 353 uIleLeuPheGluGlyValGluTyrIleAlaArgAlaGlyTrpThrProGluGlyValTy 373
Db 1261 CTCCCTTTTCCCAAGTGGAGTACATCGCCCGGCTGGGACACGGGACGGCAATA 1320
QY 373 rAlaTrpSerIleLeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSerPr 393
Db 1321 TGCTGTGGCCAGTTCCTGGACCGTCCCGAGCAACGCTTCAGCTTGTCTCTCGCCGCC 1380
QY 393 oGluLeuPheIleProValGluAspAspValMetGluArgGlnArgLeuIleGluSerVa 413
Db 1381 TGCTCTCTCTATCCGGCCGTTGAGAGTGGAGGCCCGCAGCGCAGCGCTGCCAGCGCT 1440
QY 413 lProAspSerValThrProLeuIleIleTyrGluGluThrThrAspIleTrpIleAsnI1 433
Db 1441 CCCCAGGAATGTGCAGGCCCTTTGTCTATGATGAAGAGTCAACCAATGTCTGGATCAAGT 1500
QY 433 eHisAspIlePheHisValPheProGlnSerHis---GluGluGluIleGluPheIlePh 452

Db 1501 CCAGACATCTTCCACCCGTTTCTCAGGCTGAGGCCAGCAGGAGTTTGTTCCTCTCG 1560
QY 452 eAlaSerGluCysLysThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLysG1 472
Db 1561 TGCCCAACGAATGCAAGACTGGCTTCTGCCACCTGTACAGGTTACAGGTTCACAGTAAAC 1620
QY 472 uSerLysTyrLysArgSerSerGlyGlyLeuProAlaProSerAspPheLysCysProI1 492
Db 1621 CAAGGACTATGACTGGAGCGGAACCCCTCAGCCCTCAGCAAGATGAGTTTAAAGTGCCCAT 1680
QY 492 eIyegluGluIleAlaIleThrSerGlyGluTrpGluValLeuGlyArgHisGlySerAs 512
Db 1681 CAAGGAGAGGTTCGCTTACCTGACAGTGGCGAGTGGAGGTTCTGTGAGGCGATGGCTCAA 1740
QY 512 nIleGlnValAspGluValArgArgLeuValTyrPheGluGlyThrLysAspSerProLe 532
Db 1741 GATCTGGTCAACGAGCAGCAGAGCTGGTGTACTTTCAAGGTACAAAGACACACCGCT 1800
QY 532 uGluHisIleLeuTyrValValSerTyrValAsnProGlyGluValThrArgLeuThrAs 552
Db 1801 GGAACATCACCTCTATGTGTGTACGTACAGTACAGGCGCAGATCGTGGGCTCACAC 1860
QY 552 pArgGlyTyrSerHisSerCysLysIleSerGlnHisCysAspPhePheIleSerLysTy 572
Db 1861 GCTCGGCTTCTCCACAGCTGCTCCATGAGCCAGAGCTTCGACATGTTCTGTGATCACA 1920
QY 572 rSerAsnGlnLysAsnProHisCysValSerLeuTyrLysLeuSerSerProGluAspAs 592
Db 1921 CAGCAGTGTGAGCAGCCACCTGTGTACATGTGTACAGTGTGAGCGGCCGCGATGATGA 1980
QY 592 pProThrCysLysThrLysGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuPr 612
Db 1981 CCCACTGCACAAGCAACCAACGCTTCTGGGCGCAGCATGTAGAGGAGCAATTTGCCCC 2040
QY 612 oAspTyrThrProProGluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrG1 632
Db 2041 AGACTATGTGCCCCCTGAGATCTTCCACTTCCACCCGTCAGAGCGTGAGCTCTACCG 2100
QY 632 yMetLeuTyrLysProHisAspLeuGlnProGlyLysLysTyrTrpThrValLeuPheI1 652
Db 2101 CATGATCTACAGGCCACACACCTTGCACCTGGAGGAGGAGCACCACCTGTCTTCTGT 2160
QY 652 eTyrGlyGlyProGlnValGlnLeuValAsnAsnArgPheLysGlyValLysTyrPheAr 672
Db 2161 CTATGGGGGCCCAACAGGTGAGTGTGTGTGAACACTCCTTTAAGGGCATCAATACCTGC 2220
QY 672 gLeuAsnThrLeuAlaSerLeuGlyTyrValValValValIleAspAsnArgGlySerCy 692
Db 2221 GCTAAATACACTGGCATCTTGGGCTATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2280
QY 692 shiArgGlyLeuLysPheGluGlyAlaPheLysTyrLysMetGlyGlnIleGluIleAs 712
Db 2281 TCAGCGGGGCGCTGCACTTCGAGGGGGGCGCTGAAAAAATCAAAATGGCGAGGTGAGTGA 2340
QY 712 pAspGlnValGluGlyLeuGlnTyrLeuAlaSerArgTyrAspPheIleAspLeuAspAr 732
Db 2341 GACACAGGTGGNAGGCTGACGTGAGTGGCTGAGAGATGGCTTCAATGACTTGAAGCG 2400
QY 732 gValGlyIleHisGlyTrpSerTyrGlyTyrLeuSerLeuMetAlaLeuMetGlnAr 752
Db 2401 AGTCGCCATTCATGGCTGTCTACGGCGGCTTCTCTACTCATGGGCTCATCCACAA 2460
QY 752 gSerAspIlePheArgValAlaIleAlaGlyAlaProValThrLeuTrpIlePheTyrAs 772
Db 2461 GCCACAAGTGTTCAGGTAGCCATTGGCGGCGCTCTCTCTCTCTCTCTCTCTCTCTCT 2520
QY 772 pThrGlyTyrThrGluArgTyrMetGlyHisProAspGlnAsnGluGlnGlyTyrTrLe 792
Db 2521 CACAGGTACACGGAACGATACATGATGTCCCGGAAAAATACAGCAGGAGCTATGAGGC 2580
QY 792 uGlySerValAlaMetGlnAlaGlyLysPheProSerGluProAsnArgLeuLeuLe 812


```

Qy 477 gSerSerGlyGlyLeuProAlaProSerAspPheIysCysProIleLysGluGluIleAl 497
Db 780 GAGTGAAGCCCTTCAGCCCGCGGAGAGATGAATTAAGTGGCCCAATTAAGAGAGATTGC 839
Qy 497 aileThrSerGlyGluThrGluValLeuGlyArgHisGlySerAsnIleGlnValAspGI 517
Db 840 TCTGACCAAGCGGTGATGGAGGTTTGGCGAGGCACGGCTCCCAAGATCTGGGTCAATGA 899
Qy 517 uVala-gArgLeuValTyPheGluGlyThrIysAspSerProLeuGluHisIleLeuTy 537
Db 900 GGAGACCAAGCTGGTACTTCCAGCGCACCAAGGACACCGCGCTGGAGACACACCTCTA 959
Qy 537 rValValSerTyTyValAsnProGlyGluValThrArgLeuThrAspArgGlyTySerHi 557
Db 960 CGTGGTCACTATGAGCGCGCGCGGAGATCGTACGCTCTACCGCCCGCGCTCTCCCA 1019
Qy 557 sSerCysCysIleSerGlnHisCysAspPhePheIleSerIysTySerAsnGlnLysAs 577
Db 1020 TAGCTGTCCATGAGCCAGCACTTCGACATGTTGTCAGCCACTACAGCAGCGTGAGCAC 1079
Qy 577 nProHisCysValSerLeuTyIysLeuSerSerProGluAspAspProThrCysIysTh 597
Db 1080 GCGCCCTGGTGCAGCTCTACAAAGCTGAGCGCGCCCGCAGCAGCACCCCTTCGCAAGCA 1139
Qy 597 rLysGluPheThrAlaThrIleLeuAspSerAlaGlyProLeuProAspTyThrProPr 617
Db 1140 GCCCGCTTCTGGGCTAGCATGATGAGCGAGCAGCAGCTGCCCGCCGATATGTTCTTCC 1199
Qy 617 oGluIlePheSerPheGluSerThrThrGlyPheThrLeuTyGlyMetLeuTyLysPr 637
Db 1200 AGAGATTTTCATTTCCACAGCGCTCGGATGCGGCTCTACGCGCATGATCTCAAGCC 1259
Qy 637 oHisAspLeuGlnProGlyLysTyTyProThrValLeuPheIleTyGlyGlyProGI 657
Db 1260 CCACGCTTGCAGCGAGGAGAACACACCCGCTCTCTTTGATATGAGGCGCCCA 1319
Qy 657 nValGlnLeuValAsnAsnArgPheIysGlyValIysTyTyPheArgLeuAsnThrLeuAl 677
Db 1320 GGTGAGTGTGTGAATTAACCTCTTCAAGGATCAAGTACTTGGCGCTCAACACATGGC 1379
Qy 677 aSerLeuGlyTyValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLy 697
Db 1380 CTCCCTGGGCTACCGCGTGTGTGATGAGCGAGCGGCTCTCTGTACGAGGCGCTTCG 1439
Qy 697 sPheGluGlyAlaPheIysTyTyMetGlyGlnIleGluIleAspAspGlnValGluGI 717
Db 1440 GTTCGAGGGGCCCTGAACCAATGGCCAGGTGAGATCGAGGACCGAGGTGGAGGG 1499
Qy 717 yLeuGlnTyLeuAlaSerArgTyTyPheIleAspLeuAspArgValGlyIleHisGI 737
Db 1500 CTGCACTTCTGGCGGAGAGTATGGCTTTCATCGACCTGAGCGGATGGCATCCATGG 1559
Qy 737 yTrpSerTyTyGlyTyLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheAr 757
Db 1560 CTGGTCCATCGGGGCTCTCTCTCGCTCAATGGGCTTAATCCACAGCCCGAGGTTCAA 1619
Qy 757 gValAlaIleAlaGlyAlaProValThrLeuThrPheIleTyTyAspThrGlyTyThrGI 777
Db 1620 GGTGGCCATCGGGGTCGCCCGCTCGCTCGATCGGCTGATGAGGCTACGACAGCGGTACACTGA 1679
Qy 777 uArgTyTyMetGlyHisProAspGlnAsnGluGlnGlyTyTyTyLeuGlySerValAlaMe 797
Db 1680 GCGCTACATGAGCGCTCTCGAGAACACACCGCGCTATGAGCGGGTTCGGTGGCCCT 1739
Qy 797 tGlnAlaGluIysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAs 817
Db 1740 GCAGCTGAGAGAGCTGCCCAATGAGCCCAACCGCTTCTTATCTCTCAGCGGCTTCTGGA 1799
Qy 817 pGluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyTy 837
Db 1800 CGAAACGTCGACTTTTCCACACAACTCTCTCTCTCCCACTGATCCGAGCGGAA 1859

```

```

Qy 837 sProTyAspLeuGluIleTyTy 844
Db 1860 ACCTTACCAGCTCCAGGAATAC 1881

RESULT 11
CR627380 3128 bp mRNA linear HTC 22-SEP-2004
LOCUS Homo sapiens mRNA; cDNA DKF2p762F117 (from clone DKF2p762F117).
DEFINITION CR627380
ACCESSION CR627380.1 GI:50949855
VERSION HTC.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 3128)
AUTHORS Ottenwaelder,B., Obermaier,B., Deutschenbaur,S., Schaiipp,A.,
Mewes,H.W., Weil,B., Amid,C., Osanger,A., Pobo,G., Han,M. and
Wiemann,S.
CONSTRM The German cDNA Consortium
TITLE Direct Submission
JOURNAL Submitted (22-SEP-2004) MIPS, Ingolstaedter Landstr.1, D-85764
Neuherberg, GERMANY
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; cDNA
sequenced by Medigenomix (Martinsried/Germany) within the cDNA
sequencing consortium of the German Genome Project.
This clone (DKF2p762F117) is available at the RZPD Deutsches
Resourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.
Please contact RZPD for ordering:
http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKF2p762F117
Further information about the clone and the sequencing project is
available at http://mips.gsf.de/projects/cdna/.

FEATURES
Location/Qualifiers
1..3128
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKF2p762F117"
/tissue_type="melanoma (MeWo cell line)"
/clone_lib="762 (synonym: hm12). Vector pSport1; host
DH10B; sites NotI + SalI"
/dev_stage="adult"
/note="dipeptidylpeptidase 9, N-terminus truncated"

gene
1..3128
/gene="DKF2p762F117"
<12..1712
/gene="DKF2p762F117"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAH10477.1"
/db_xref="GI:50949856"
/translation="RKTDVRYPRVTGSKNPKIALKLAERFQDSQGIKIVSTQKELVOP
FSSLPFKVYIARAGWTRDQKIYAWFLDRPQWLQVLPPALFTIPSTENSEORLAS
ARAVRNQPVYVYEVTVNINVDIIFYFPOSEGEDELCLFRANECKTGCHLYKV
TAVLKSDYDSESPGDEFKCPTKEIALTSGEWEVLARHSGKIVWNEETKLYVF
QGTQDTPLHLLYVVSVEAAGEIVRLTTPGFSHSCMSQNFDMFVSHSVSTPCVH
VYKLSGDDPDLHKQPRFASMMEAASCDDYVPEIHFPHTRSDVRLYGMVYKPHAL
QPKRGHFTLVYGGPQVQVWNSFKGIKYLRLNTLASLGYAVVVDGRGSCORGLRF
EGALKNQMGVEIIDEVGEQFVAEKYGFIDLSRVALHGSVYGGFSLMGLIHKQVLF
KVATAGAPVTVMAYDTGYRYNDVPENNQHYEAGSVLHVKEKLPNPNRLLILHG
FLDENVHFEHTNVLVSLIRAGKPYQIYIPNHRHSIRCPESGSEHYEVTLHLLQLEYL
"

Alignment Scores:
Pred. No.: 2,91e-216 Length: 3128
Score: 1991.50 Matches: 359
Percent Similarity: 78.4% Conservative: 85
Best Local Similarity: 63.4% Mismatches: 121
Query Match: 42.4% Indels: 1
Gaps: 1
DB: 4

```

US-10-825-632-1 (1-882) x CR627380 (1-3128)

QY 308 ArgArgAlaAspSerPheArgTyrProLysThrGlyThrAlaAsnProLysValThrPhe 327
 DB 12 AGAAGACGGACTCGTATCGGTACCCAGGACAGGCGACGAGAAATCCCAAGATTGCCCTTG 71
 QY 328 LysMetSerGluLeuMetIleAspAlaGluGlyArgIleAspValIleAspLysGlu 347
 DB 72 AAATCGCTGAGTTCCAGACTGACAGCCAGGCGAGATCGTCTCGACCCAGGAGAGAG 131
 QY 348 LeuLeuGlnProPheGluLeuLeuPheGluGlyValGluTyrIleAlaArgAlaGlyTyr 367
 DB 132 CTGGTGCAGCCCTTACGCTCGCTGTTCCCGAAGGTGGAGTACATCGCCAGGCGCGGTGG 191
 QY 368 ThrProGluGlyLysTyrAlaTyrSerIleLeuLeuAspArgSerGlnThrArgLeuGln 387
 DB 192 ACCCGGATGGCAATACGCTCGGCGCATGTTCTCGACCGCGCCCGAGTGGTCCAG 251
 QY 388 IleValLeuLeuSerProGluLeuLeuPheIleProValGluAspValMetGluArgGln 407
 DB 252 CTCGTCTCTCCCGCGCCCTGTTTCATCCCGGACACAGAGAAATGAGGACGCGGCTA 311
 QY 408 ArgLeuIleGluSerValProAspSerValThrProLeuIleIleTyrGluGluThrThr 427
 DB 312 GCCTCTCCAGAGCTGTCCTCCAGGAATGTCACCGCTATGTGGTGTACGAGGAGGTCA 371
 QY 428 AspiletrpileAsnIleHisAspIlePheHisValPheProGlnSerHis---GluGlu 446
 DB 372 AACGTGTGATCAATGTTCATGACATCTTCTATCCCTTCCCAATACAGGGAGAGAC 431
 QY 447 GluLeuGluPheIlePheAlaSerGluCysLysThrGlyPheArgHisLeuTyrIle 466
 DB 432 GAGCTGTCTTCTCCCGCCCAATGAATGACAGCGGCTTCTGCCATTTGTACAAAGTC 491
 QY 467 ThrSerIleLysGluSerLysTyrLysArgSerSerGlyLeuProAlaProSer 486
 DB 492 ACCCGCTTTAAATCCAGGCTACGATTGGAGTGAGCCCTTCAGCCCGCGGGAAGAT 551
 QY 487 AspPheLysCysProLysGluGluIleAlaIleThrSerGlyGluTyrGluValLeu 506
 DB 552 GAATTTAAGTGCCCCATTAAAGGAAGAGATTGCTCTGACACGCGGTGAATGGAGGTTTG 611
 QY 507 GlyArgHisGlySerAsnIleGlnValAspGluValArgArgLeuValTyrPheGluGly 526
 DB 612 GCGAGGCACGGCTCCAAGATCTGGGTCAATGAGGAGACCAAGCTGGTGTACTCTCAGGGC 671
 QY 527 ThrLysAspSerProLeuGluHisHisLeuTyrValValSerTyrValAsnProGlyGlu 546
 DB 672 ACCAAGACACGCGCTGGAGCACCACTCTAGTGGTACGCTATGAGCGCGCGCGGAG 731
 QY 547 ValThrArgLeuThrAspArgGlyTyrSerHisSerCysCysIleSerGlnHisCysAsp 566
 DB 732 ATCGTAGCCTCACACGCGCGGCTTCTCCATAGTGTCTCATGAGCCAGCAATCTGCAC 791
 QY 567 PhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeuTyrLysLeu 586
 DB 792 ATGTTCGTACACACTACAGCGCTGAGCAGCGCCCTCGTGGTGCAGCTCTACAGCTG 851
 QY 587 SerSerProGluAspAspProThrCysLysThrLysGluPheTyrAlaThrIleLeuAsp 606
 DB 852 AGCGCGCCGACGACGACCCCTTGCAAGCAGCGCCCGCTTCTGGGTAGCATGATGAG 911
 QY 607 SerAlaGlyProLeuProAspTyrThrProProGluIlePheSerPheGluSerThrThr 626
 DB 912 GCAGCCAGCTGCCCCCGGATATGTTCTCTCAGAGATCTTCATTTCCACACCGCGCTG 971
 QY 627 GlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGlyLysLysTyr 646
 DB 972 GATGTGGGCTCTACGCGCATGATCTCAAGCCGCCAGCGCTTGCACCCAGGAGGAAGAC 1031
 QY 647 ProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsnAsnArgPheLys 666

DB 1032 CCCACCGTCTCTTTGTATATATGAGGGCCCCCAGGTGCAGCTGGTGAATAACTCTCTCAA 1091
 QY 667 GlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValValValValle 686
 DB 1092 GGCATCAAGTACTTTCGGGCTCAACACACTGGCTCTCCCTGAGCTACGCCGTGGTGTGATT 1151
 QY 687 AspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLysTyrLysMet 706
 DB 1152 GACGCGAGGGCTCTCTGACAGAGGGCTTCGTTTGGAGGGCCCTGAAAAACCAATG 1211
 QY 707 GlyGlnIleGluLeuAspAspGlnValGluGlyLeuGlnTyrLeuAlaSerArgTyrAsp 726
 DB 1212 GGCACAGTGGAGATCGAGGACCAAGTGGAGGGCTCGAGTTCGTGGCCGAGAAATATGC 1271
 QY 727 PheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyrGlyGlyTyrLeuSerLeu 746
 DB 1272 TTCATCGACCTGAGCCGAGTTGCCATCCATGGCTGGTCTTACGGGGCTTCTCTCGCTC 1331
 QY 747 MetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAlaProValThr 766
 DB 1332 ATGGGGCTAATCCACAAGCCCGAGGTGTTCAAGTGGCCATCGCGGTGCCCGGTCA 1391
 QY 767 LeuTyrIlePheTyrAspThrGlyTyrThrGluArgTyrMetGlyHisProAspGlnAsn 786
 DB 1392 GTCTGATGGCTTACGACACAGGCTACACTGAGGGCTACATGAGCGTCTCTGAGAACA 1451
 QY 787 GluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGluLysPheProSerGluPro 806
 DB 1452 CAGCAGGCTATGAGCGGGTTTCGTGGCCCTCCACGTGGAGAGAGCTGCCAATGAGCC 1511
 QY 807 AsnArgLeuLeuLeuHisGlyPheLeuAspGlnValHisPheAlaHisThrSer 826
 DB 1512 AACCGTTCGTATCTCCACGGCTTCTGGAGAAACGTGCACATTTTCCACACAAC 1571
 QY 827 IleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAspLeuGlnIleTyrProGln 846
 DB 1572 TTCCTGCTCCCACTGATCCGAGCAGGAGAAACCTTACAGCTCCAGATCTACCCCAAC 1631
 QY 847 GluArgHisSerIleArgValProGluSerGlyGluHisTyrGluLeuHisLeuLeuHis 866
 DB 1632 GAGAGACACAGTATTCGCTGCCGAGTGGGGCAGCACTATGAAGTCACTGCTGTGCAC 1691
 QY 867 TyrLeuGlnGluAsnLeu 872
 DB 1692 TTTCTACAGGAATACCTC 1709

RESULT 12
 DQ053109
 LOCUS 2767 bp DNA linear GSS 02-JUN-2005
 DEFINITION Homo sapiens DPP9 gene, VIRTUAL TRANSCRIPT, partial sequence,
 genomic survey sequence.
 ACCESSION DQ053109
 VERSION DQ053109.1 GI:66899056
 KEYWORDS GSS.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Sukariyota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.
 REFERENCE 1 (bases 1 to 2767)
 AUTHORS Nielsen,R., Bustamante,C., Clark,A.G., Gnanowski,S., Sackton,T.B.,
 Hubisz,M.J., Fiedel-Alon,A., Tanenbaum,D.M., Civelio,D.,
 White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
 TITLE A Scan for Positively Selected Genes in the Genomes of Humans and
 Chimpanzees
 JOURNAL (ex) Plos Biol. 3 (6), E170 (2005)
 PUBMED 15869325
 REFERENCE 2 (bases 1 to 2767)
 AUTHORS Nielsen,R., Bustamante,C., Clark,A.G., Gnanowski,S., Sackton,T.B.,
 Hubisz,M.J., Fiedel-Alon,A., Tanenbaum,D.M., Civelio,D.,
 White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
 TITLE Direct Submission
 JOURNAL Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive,

Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering them based on alignment. Translation starts at the beginning of alignment.

FEATURES
source Location/Qualifiers
1. 2767
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="19"
1. 2767
/gene="DPP9"
/locus_tag="HC13651"

ORIGIN
Alignment Scores:
Pred. No.: 6,47e-188 Length: 2767
Score: 1744.00 Matches: 362
Percent Similarity: 53.8% Conservative: 90
Best Local Similarity: 43.1% Mismatches: 386
Query Match: 37.1% Indels: 2
DB: 11 Gaps: 2

US-10-825-632-1 (1-882) x DQ053109 (1-2767)

QY 35 PheTyrValGluArgTyrSerTrpSerGlnLeuLysLeuLeuAlaAspThrArgLys 54
DB 245 TTCAGGTGCAGAGCACTGTGGACGGCTCCGAGCATCATCCAGCGCGCAAG 304
QY 55 TyrHisGlyTyrMetMetAlaLysAlaProHisAspPheMetPheValLysArgAsnAsp 74
DB 305 TACTCGGGCCCTATTGTCAACAGGGCCGCCACGACTTCAGTTTGTGCAGAGCGGAT 364
QY 75 ProAspGlyProHisSerAspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGlu 94
DB 365 GAGTCTGGGGCCCTCCACCGCTCTACTACCTGGGAATGCCATGTGGACGGCGAG 424
QY 95 AsnThrLeuPheTyrSerGluIleProLysThrIleAsnArgAlaAlaValLeuMetLeu 114
DB 425 AACTCCCTCTCTACTCTGAGATTCCCAAGAGTCCGGAAGAGGCTCTGTCTCTCG 484
QY 115 SerTrpLysProLeuLeuAspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArg 134
DB 485 TCCTGGAGCAGAGTCTGGATCATTTCCAGCGCACGCCACCCATGGGGTCTACTCTCG 544
QY 135 GluGluGluLeuLeuArgGluArgLysArgIleGlyThrValGlyIleAlaSerTyrAsp 154
DB 545 GAGGAGGAGCTGCTGAGGAGCGGAAACGCTGGGGGTCTTCGGCATCACCTCTTACGAC 604
QY 155 TyrHisGlnGlySerGlyThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLys 174
DB 605 TTCCACACGAGAGTGGCTCTCTCTTCAGCGCAGCAGCAGCTCTTCCACCTGCGCGC 664
QY 175 AspGlyGlyProGlnGlyPheThrGlnGlnProLeuArgProAsnLeuValGluThrSer 194
DB 665 GACGCGGCAAGACGGCTTCATGGTCTCCCTATGAACCGCTGGAAATCAAGACCCAG 724
QY 195 CysProAsnIleArgMetAspProLysLeuCysProAlaAspProAspTrpIleAlaPhe 214
DB 725 TGCTCAGGCGCCCGGATGGACCCCAAAATCTGCCATCTGCCGACCTCTCTCTNNNNN 784
QY 215 IleHisSerAsnAspIleTrpIleSerAsnIleValThrArgGluGluArgLeuThr 234
DB 785 NNN 844
QY 235 TyrValHisAsnGluLeuAlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThr 254
DB 845 NNN 904
QY 255 PheValLeuGlnGluPheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGlu 274
DB 905 NNN 964

QY 275 ThrThrProSerGlyGly---LysIleLeuArgIleLeuTyrGluGluAsnAspGluSer 293
DB 965 NNNNNNGTTTCAGAGGGCCCTCAAGACGCTCGCAATCCTGTATGAGAGAGTCGATGAGTCC 1024
QY 294 GluValGluIleIleHisValThrSerProMetLeuGluThrArgArgAlaAspSerPhe 313
DB 1025 GAGGTGGAGGTTCATTCACGTCCTCTCTCGCTAGAGAAGGAAGGACGACCTCGTAT 1084
QY 314 ArgTyrProLysThrGlyThrAlaAsnProLysValThrPheLysMetSerGluIleMet 333
DB 1085 CGGTACCCAGGACAGGACGAGCAAGATTCACGATTCCTTGAATTCGCTGAGTTCACG 1144
QY 334 IleAspAlaGluGlyArgIleIleAspValIleAspLysGluLeuIleGlnProPheGlu 353
DB 1145 ACTGACAGCCAGGCAAGATCGTCTGACCCAGAGAGGAGCTGTGCGAGCCCTTCAGC 1204
QY 354 IleLeuPheGluGlyValGluTyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyr 373
DB 1205 TCGCTGTTCCGAAGGTGGAGTACATCGCCAGGCGCGGTGGACCCGGGATGCAATAC 1264
QY 374 AlaTrpSerIleLeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSerPro 393
DB 1265 GCCTGGCCCATGTTCTTGGACCGGCCCCAGCAGTGGCTCCAGCTCTCTCTCTCCCGG 1324
QY 394 GluLeuPheIleProValGluAspValMetGluArgGlnArgLeuIleGluSerVal 413
DB 1325 GCCCTGTTTCATCCGAGCAGACAGAGATGAGGAGCAGCGCTAGCCTCTCCAGAGCTGC 1384
QY 414 ProAspSerValThrProLeuIleIleTyrGluThrThrAspIleTrpIleAsnIle 433
DB 1385 CCCAGATGTCAGCGCTATGTGGTGTACGAGGAGGTACCAACCTCTGGATCAATGTT 1444
QY 434 HisAspIlePheHisValPheProGlnSerHis---GluGluGluIleGluPheIlePhe 452
DB 1445 CATGACATCTTCTATCTCCCTTCCCAATCAGAGGAGAGGAGCAGCTCTCTCTCTCCG 1504
QY 453 AlaSerGluCysLysThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLysGlu 472
DB 1505 GCCAATGAATCAAGACCGGCTTCTGCCATTTGTACAAAGTACCCCGCTTTTAAATCC 1564
QY 473 SerLysTyrLysArgSerSerGlyLeuProAlaProSerAspPheLysCysProIle 492
DB 1565 CAGGGCTAGATTGAGTGGATGACCTTCAGCCCGGGGAGAGATGAATTAAGTCCCAAT 1624
QY 493 LysGluGluIleAlaIleThrSerGlyGluTrpGluValLeuGlyArgHisGlySerAsn 512
DB 1625 AAGGAAGAGATTGCTCTGACCGCGGTGAATGGAGGTTTTGGCGAGGACGCGCTCCAAG 1684
QY 513 IleGlnValAspGluValArgArgLeuValTyrPheGluGlyThrLysAspSerProLeu 532
DB 1685 ATCTGGGTCAATGAGGAGACCAAGCTGTGTACTTCCAGGGGCCAAGGACGCGCGCTG 1744
QY 533 GluHisIleLeuTyrValValSerTyrValAsnProGlyGluValThrArgLeuThrAsp 552
DB 1745 GAGCACCACTCTAGTGGTCAAGTATGAGCGCGCGCGGAGATCGTACGCTCACCACG 1804
QY 553 ArgGlyTyrSerHisSerCysIleSerGlnHisCysAspPhePheIleSerLysTyr 572
DB 1805 CCCGGCTTCTCCATAGCTGTCTCATGAGCCAGNNNNNNNNNNNNNNNNNNNNNNNN 1864
QY 573 SerAsnGlnLysAsnProHisCysValSerLeuTyrLysLeuSerSerProGluAspAsp 592
DB 1865 NNN 1924
QY 593 ProThrCysLysThrLysGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuPro 612
DB 1925 NNN 1984
QY 613 AspTyrThrProProGluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGly 632
DB 1985 NNN 2044
QY 633 MetLeuTyrLysProHisAspLeuGlnProGlyLysLysTyrProThrValLeuPheIle 652


```

QY 804 rGluProAsnArgLeuLeuLeuHis sGlyPheLeuAspGluAsnValHisPheAlaHi 824
DB 759 TGAACCAAACTGTTACTGCTCTTACATGTTTCTCGGATGAGAATGCCATTTTGACA 818
QY 824 sThrSerLeuLeuSerPheLeuValArgAlaGlyLysProTyrAspLeuGlnIleTy 844
DB 819 TACAGTATATTACTGAGTTTTTTAGTGAGGCTGGAGCCATATGATTACAGATCTA 878
QY 844 rProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyrGluLeuHisLe 864
DB 879 TCCTCAGAGAGACACAGCAAGAGTTCTGAATCGGAGAACATTATGAATGCACTCT 938
QY 864 uLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeuLysValIle 882
DB 939 TTTGCACTACCTTCAAGAAAACCTTGGATCACGTATTGCTGCTTAAAGTGATA 993

RESULT 14
AF175225 1292 bp mRNA linear HTC 01-AUG-2003
LOCUS Homo sapiens tissue-type aorta MSTP135 mRNA, complete cds.
ACCESSION AF175225
VERSION AF175225.1 GI:33338055
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1 (bases 1 to 1292)
AUTHORS Zhao,B., Xu,H.S., Tong,Y.K., Sheng,H., Qin,B.M., Liu,Y.Q., Liu,B.,
Wang,X.Y., Zhang,Q., Song,L., Gao,Y., Zhang,C.L., Ye,J., Ji,X.J.,
Liu,B.H., Lu,H., Chen,J.Z., Cai,M.Q., Zheng,W.Y., Teng,C.Y.,
Liu,Q., Yu,L.T., Lin,J., Gong,Q., Zhang,A.M., Gao,R.L. and Hui,R.T.
Direct Submission
JOURNAL Submitted (04-AUG-1999) Molecular Medicine Center for
Cardiovascular Disease, Cardiovascular Institute, CAMS & PUMC, 167,
Bei Li Shi Lu, Beijing, 100037, P.R. China
FEATURES
source Location/Qualifiers
1..1292
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="aorta"
142..1023
/codon_start=1
/product="MSTP135"
/db_xref="GI:33338056"
/translation="MTOLAKQRNFPFPFIQVLFITLILQKFLKLVLLDLHCWGCS
TSPHSGTGLERNILLXLFYGGPKVOLVNNRFGKVFRLNLPFLGVVAVXGIDNRGS
CPRLKPEGAFKPMGQIEIDDOVGLGYLASHYDFIDLDVRGHWGYSYGLISML
MQRSDIFRVALAGAPVTLNIFPYGTGYRYMGPDQNEQGYLGSVMAQAEKPESEPN
RLLLHGFLENVHFATSIILLSFLVRAGKPYDLQIYPOERHSIRVPSEGEHYELHLL
HYLOENLGSRIAALKVT"
CDS
Alignment Scores:
Pred. No.: 9,31e-186 Length: 1292
Score: 1720.50 Matches: 325
Percent Similarity: 97.1% Conservative: 5
Best Local Similarity: 95.6% Mismatches: 8
Query Match: 36.6% Indels: 2
DB: 4 Gaps: 1

US-10-825-632-1 (1-882) x AF175225 (1-1292)

QY 545 GlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCysIleSerGlnHis 564
DB 3 GGAGAGGTGACAAAGCTGACTGACCGTGGCTACTCACATTTCTTGCTCATGTCAGTCAC 62
QY 565 CysAspPhePheIleSerIleTyrSerAsnGlnLysAsnProHisCysValSerLeuTyr 584

```

```

DB 63 TGTGACTTCTTTATAGTAAGTATAGTAACCAAGAAATCCACACTGTGTGCTCCCTTTAC 122
QY LysLeuSerSerProGluAspAspProThrCysLeuThrLysGluPheThrAlaThrIle 604
DB 123 AAGCTATCAGTCTCTGAAGATGACCCCACTTGCAAAACAAGGAATTTTGGGCCACATT 182
QY LeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePheSerPheGluSer 624
DB 183 TTGGATTCCAGCAGGTCTCTCTCTGACTATCTCTCCAGAAATTTCTCTTTGAAAGT 242
QY ThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu---GlnProGly 643
DB 243 ACTACTGGAATTACATTTGTATGGGATGCTCTACAAAGCCCTCATGGAGCTACAGGCCTGA 302
QY LysLysTyrProThr-ValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsnAs 663
DB 303 AAGAAATATCTCTACTGNGGCTGTTTCATATATGGGGGCCCTTAAGGTGCAGTTGTGANTAA 362
QY nArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVa 683
DB 363 TCGGTTTAAAGGAGTCAAGTTTTTCCCTTGATATACCTTACCTTTTATAGTTATGTGT 422
QY lValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheTy 703
DB 423 TGNAGGGATAGACACACAGGGGATCTCTCCCCGAGGGCTTAAATTTGAAGGGCCCTTAA 482
QY sTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyrLeuAlaSe 723
DB 483 ATTTAAATGGTCAATAGAAATTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTC 542
QY rArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyrGlyTy 743
DB 543 TCGATATGATTTCATTGACTTCGTGTGGGCATCCACGCTGTCTCTATGGAGATA 602
QY rIleuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAl 763
DB 603 CCTCTCCCTGATGGCATTAATGCAGAGGTGAGATATCTTCAGGGTGTCTATTGCTGGGC 662
QY aProValThrLeuTyrPheTyrAspThrGlyTyrThrGluArgTyrMetGlyHisPr 783
DB 663 CCCAGTCACTCTGTGGATCTTCTATGATACAGGATACCGGAACGTTATATGGTCAACC 722
QY oAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGluLysPhePr 803
DB 723 TGACCAAGATGACACAGGCTATTACTTAGGATCTGTGCCCATGCAAGCAGAAAGTTCCC 782
QY oSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnValHisPheAl 823
DB 783 CTCTGACCAATCGTTTACTGCTCTTACATGTTTCTCGATGAGAATGTCCATTTTGC 842
QY aHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAspLeuGlnI 843
DB 843 ACATACCAGTATATTACTGAGTTTTTTAGTGAGGGCTGGAAGCCATATGATTTACAGAT 902
QY eTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyrGluLeuHi 863
DB 903 CTATCTCAGAGAGACACACAGCAAGAGTTCTCGAATCGGGAGAACATTATGAACCTGCA 962
QY sLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLysValIle 882
DB 963 TCTTTTGCACCTACCTTCAAGAAAACCTTGGATCAGCTATTGCTGCTCTTAAAGTGATA 1020

RESULT 15
BM557438 1042 bp mRNA linear EST 20-FEB-2002
LOCUS AGENCOURT_6578992 NIH_MGC_41 Homo sapiens cdna clone IMAGE:5466759
5', mRNA sequence.
ACCESSION BM557438
VERSION BM557438.1 GI:18799430
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae, Homo.

REFERENCE

1. (bases 1 to 1042)
NIH-MGC <http://mgs.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTF
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L1CM1967 row: k column: 16
High quality sequence stop: 697.
Location/Qualifiers
1..1042
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clones IMAGE:3466759"
/issue_types="amelanotic melanoma, cell line"
/lab_hosts="NIH_MGC_41"
/note="Organ: Skin; Vector: pOT7; Site: 1: XhoI; Site: 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

FEATURES

source
1..1042
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clones IMAGE:3466759"
/issue_types="amelanotic melanoma, cell line"
/lab_hosts="NIH_MGC_41"
/note="Organ: Skin; Vector: pOT7; Site: 1: XhoI; Site: 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

ORIGIN

Alignment Scores:
Pred. No.: 8.27e-183 Length: 1042
Score: 1693.50 Matches: 324
Percent Similarity: 94.0% Conservative: 3
Best Local Similarity: 93.1% Mismatches: 15
Query Match: 36.0% Indels: 6
DB: 3 Gaps: 1
US-10-825-632-1 (1-882) x BM557438 (1-1042)

QY 504 GluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArgLeuValTyr 523
DB 1 GAAGTTCTTGGCCGGCATGGATCTATATCCAAAGTTGATGAAGTCAGAAAGCTGGTATAT 60
QY 524 PheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSerTyrValAsn 543
DB 61 TTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTAGTCAGTACGTACGTAAAT 120
QY 544 ProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysIleSerGln 563
DB 121 CCTGGAGAGGTGACAAAGGCTGACTGACCGGTGGCTACTCATCTTGTGTCATCAGTCAG 180
QY 564 HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeu 583
DB 181 CACTGTGACTCTTTATTAAGTAAGTATAGTAACCAAGAAATCCACACTGTGTGCCCTT 240
QY 584 TyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThr 603
DB 241 TACAAGCTATCAAGTCTCTGAAGATGACCCCAACTTGCAAAACAAAGAAATTTGGGCCACC 300
QY 604 IleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePheSerPheGlu 623
DB 301 ATTTGGATTACAGAGGTCTCTTCTGACTACTACTCTCCAGAAATTTCTCTTTTGA 360
QY 624 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 643
DB 361 AGTACTACTGGATTTCATTGTATGGGATGCTCTACAAAGCCCTCATGATCTACAGCCTGGA 420

QY 644 LysLysTyrProThrValLeuPheIleTyrGlyProGlnValGlnLeuValAsnAen 663
DB 421 AAGAAATATCTACTGTGCTGTTTCATATATGGTGGTCTTCAGGTGCAGTTCGTGAATAT 480
QY 664 ArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrValVal 683
DB 481 CGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTAGCCTCTCTAGGTTATGTGTT 540
QY 684 ValValLeuAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLys 703
DB 541 GTAGTGATAGACACAGGGATCTCTGTCCAGAGGCTTAAATTTGAAGCCGCTTTANA 600
QY 704 TyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyrLeuAlaSer 723
DB 601 TATAAATGGGTCAATAGAAATTGAGATCAGGTGGAAGGACTCCAATATCTAGCTTCT 660
QY 724 ArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrPheSerTyrGlyGlyTyr 743
DB 661 CGATATGATTTTCAATTGACTTGTAGATCGTGTGGGCTCCAGGCTGGTCTTATGAGGATAC 720
QY 744 LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla 763
DB 721 CTCTCCCTGATGGCATTAATGCAGAGGTGAGATATCTTCAGGGTTGCTATTGCTGGGGCC 780
QY 764 ProValThrLeuTyrPheTyrAspThrGlyTyrThrGluArgTyrMetGlyHisPro 783
DB 781 CAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGTACGTTATATGGTCACTT 840
QY 784 AspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGluLysPhePro 803
DB 841 GACCAAGATGAACAGGGCTATTACTTANGATCTGTGGCCATGCAAGCAGAAAGTCCCCC 900
QY 804 SerGluProAsnArgLeuLeuLeuHisGly-PheLeuAspGluAsnVal-HisPheA 823
DB 901 TCTGAACCAATCGATTACTGCTTACATGCTTTTCTGGATGAGAATGTCCCATTTG 960
QY 823 IaHisThrSer-IleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAspLeu--- 841
DB 961 CACATACCCAGGATATTAATCTGGAGTTTTTTTTTNTGGAGGGGCTGAAAAAGCCCATATGGAT 1020
QY 842 -----GlnIleTyrPro 845
DB 1021 TTTACACGAATCTATCCT 1038

Search completed: April 15, 2006, 05:59:28
Job time : 8374 secs

GenCore version 5.1.7
Copyright (C) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 15, 2006, 00:36:45 ; Search time 379 Seconds
(without alignments)
4136.700 Million cell updates/sec

Title: US-10-825-632-1
Perfect score: 4700
Sequence: 1 MAAMETEQLGVIFETADC.....HLLHYLQNLGSRALAKVI 892

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 1303057 seqs, 888780828 residues
Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame_plus_p2n.model -DEV=xlip
-Q=/abs/ABSSWEB_spool/US10825632/runat_14042006_092603_4744/app_query.fasta_1
-DB=Issued Patents NA -QFWT=fastap -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFWT=pco -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abs05p -USER=US10825632 @CGN_1_1_290 @runat_14042006_092603_4744
-NCPU=6 -ICPU=3 -NO WMAP -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/1 COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5 COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/1/ina/H COMB.seq.*
6: /cgn2_6/ptodata/1/ina/PCUS COMB.seq.*
7: /cgn2_6/ptodata/1/ina/PP COMB.seq.*
8: /cgn2_6/ptodata/1/ina/RE COMB.seq.*
9: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	4700	100.0	2671	3	US-09-976-674-2
2	4700	100.0	3120	3	US-10-070-464-2
3	4680	99.6	4829	3	US-09-976-674-12
4	4443.5	94.5	2797	3	US-09-976-594-1103
5	4385.5	93.3	4685	3	US-09-976-674-22
6	4385	93.1	4676	3	US-09-976-674-20
7	4092.5	87.1	4523	3	US-09-976-674-8
8	3661.5	77.9	4309	3	US-09-976-674-14
9	2870	61.1	2617	3	US-09-976-674-4

10	2870	61.1	4219	3	US-09-976-674-28	Sequence 28, Appl	
11	2870	61.1	4302	3	US-09-976-674-24	Sequence 24, Appl	
12	2820.5	60.0	4180	3	US-09-976-674-36	Sequence 36, Appl	
13	2820.5	60.0	4263	3	US-09-976-674-34	Sequence 34, Appl	
14	2649	56.4	4076	3	US-09-976-674-32	Sequence 32, Appl	
15	2649	56.4	4159	3	US-09-976-674-30	Sequence 30, Appl	
16	2599.5	55.3	4037	3	US-09-976-674-40	Sequence 40, Appl	
17	2599.5	55.3	4120	3	US-09-976-674-38	Sequence 38, Appl	
18	2422	51.5	1669	3	US-10-070-464-6	Sequence 6, Appl	
19	1836.5	39.1	1083	3	US-10-070-464-8	Sequence 8, Appl	
20	1645.5	35.0	1197	3	US-10-070-464-4	Sequence 4, Appl	
21	1391	29.6	2411	3	US-09-976-674-26	Sequence 26, Appl	
22	1278	27.2	1356	3	US-09-976-674-10	Sequence 10, Appl	
23	1026.5	21.8	832	3	US-09-976-674-18	Sequence 18, Appl	
24	755.5	16.1	823	3	US-09-280-116-171	Sequence 171, Appl	
C	25	719.5	15.3	14555	3	US-09-902-540-1096	Sequence 1096, Ap
26	719	15.3	2946	3	US-09-902-540-9675	Sequence 9675, Ap	
27	597.5	12.7	4280	3	US-09-079-592-1	Sequence 1, Appl	
28	597.5	12.7	5496	3	US-09-462-284-1	Sequence 1, Appl	
29	572	12.2	2199	3	US-09-518-550-38	Sequence 38, Appl	
30	529	11.3	3407	3	US-10-002-593-5	Sequence 5, Appl	
31	529	11.3	3407	3	US-09-949-016-275	Sequence 275, App	
32	529	11.2	2924	3	US-10-423-714-5	Sequence 5, Appl	
33	525	11.2	2924	3	US-09-902-655-1026	Sequence 1026, Ap	
34	525	11.2	2924	6	PCT-US93-07923-1	Sequence 1, Appl	
35	525	11.2	3419	3	US-09-949-016-4579	Sequence 4579, Ap	
36	500.5	10.6	20740	3	US-09-902-540-1223	Sequence 1223, Ap	
37	498.5	10.6	2115	3	US-09-902-540-4518	Sequence 4518, Ap	
38	488.5	10.4	2214	3	US-09-902-540-4124	Sequence 4124, Ap	
39	488.5	10.4	23091	3	US-09-902-540-1204	Sequence 1204, Ap	
C	40	464	9.9	620	3	US-09-976-674-16	Sequence 16, Appl
41	461	9.8	2576	3	US-10-363-937-21	Sequence 21, Appl	
42	461	9.8	2583	3	US-09-976-674-6	Sequence 6, Appl	
43	461	9.8	4541	3	US-09-976-674-42	Sequence 42, Appl	
44	454	9.7	2815	2	US-08-230-491A-1	Sequence 1, Appl	
45	454	9.7	2815	2	US-08-619-280A-1	Sequence 1, Appl	

ALIGNMENTS

RESULT 1
US-09-976-674-2
; Sequence 2, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; PRIOR FILING DATE: 2001-10-12
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 2671
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-976-674-2

Alignment Scores:
Pred. No.: 0
Score: 4700.00
Percent Similarity: 100.0%
Best Local Similarity: 100.0%
Query Match: 100.0%
DB: 3
Length: 2671
Matches: 882
Conservative: 0
Mismatches: 0
Indels: 0
Gaps: 0

US-10-825-632-1 (1-882) x US-09-976-674-2 (1-2671)

QY 1 MetAlaAlaMetGluThrGluInLeuGlyValGluIlePheGluThrAlaAspCys 20
Db 8 ATGGCAGCAGCAATGGAACACAGACACCTGGGTGTTGAGATATTTGAACATCGCGACTGT 67
QY 21 GluGluAenIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
Db 68 GAGGAGAAATATTTGAATCACAGATCGGCCTAAATTTGAGCCCTTTATTTGTTGCGGTAT 127
QY 41 SerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
Db 128 TCCTGGAGTCAGCTTAAAGAGCTCTTTGCCGATACAGAAAATATCATGGCTACATGATG 187
QY 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
Db 188 GCTAAGGACACCATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTTCA 247
QY 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
Db 248 GACAGAACTTATACCTTGCCATGCTCTGGTGAGAACAGAGAAAATACACTGTTTATTCT 307
QY 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
Db 308 GAAATTTCCAAAAATATCAATAGACAGCAGTCTTAATGCTCTCTTTGGAAGCCTCTTTTG 367
QY 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg 140
Db 368 GATCTTTTTCAGGCAACACCTGACATATGGAATGTATTTCTCGAGNAGNAGAACTATTAGA 427
QY 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
Db 428 GNAAGNAAACGCAATGGNACNGTCGNAATTTGCTTTACGATTAATCACCCNAGGAGTGA 487
QY 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
Db 488 ACATTTCTGTTTCAAGCCGGTAGTGGAAATTTATCAGCTAAAGATGGAGGGCCACAAGGA 547
QY 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
Db 548 TTATCGCAACCACTTTAAGGCCCAATCTAGTGGAACTAGTTGTCCCAACATACGGATG 607
QY 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
Db 608 GATCCAAAATTTATGCCCTGTGATCCAGACTGGATGCTTTATATACATAGCAACGATATT 667
QY 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu 240
Db 668 TGGATATCTAACATCGTAACCCAGAGAGAAAGAGACTCACTTATGTGCACAATGAGCTA 727
QY 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu 260
Db 728 GCCACATGGAAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCCAAGAGAA 787
QY 261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280
Db 788 TTTGATAGATATTTCTGGCTATTGGTGTGTCCAAAGCTGAAACACTCCCAAGTGGTGT 847
QY 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300
Db 848 AAAATCTTTAGAAATCTATATGAAGAAAATGATGAATCTGAGGTGGAAATTAATCATGTT 907
QY 301 ThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThr 320
Db 908 ACATCCCTATGTTTGGAAACAGAGGGCAGATTCATTCGTTATCTCTTAAACAGGTACA 967
QY 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
Db 968 GCAATCTCTAAAGTCACCTTTTAAGATGTCAGAAATTAATGATGTCGAAAGAGGATC 1027
QY 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
Db 1028 ATAGATGTCATAGATAAGGAACATAATTCACACCTTTTGGAGATCTATTTTGAAGGAGTTGAA 1087
QY 361 TyrIleAlaArgAlaGlyTyrThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380

Db 1088 TATATTGCCAGAGCTGGATGGACTCCTGAGGGAAAATATGCTTGGTCCATCTCTACTAGAT 1147
QY 381 ArgSerGlnThrArgLeuGluIleValLeuIleSerProGluLeuPheIleProValGlu 400
Db 1148 CGCTCCAGACTCGCCTGCAGATAGTGTGATCTCACCTGCAATATTTATCCAGTAGAA 1207
QY 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
Db 1208 GATGATGTTATGGAAGGCAGAGACTCATGTAGCTAGTGCCTGATTTCTGAGCCACACTA 1267
QY 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
Db 1268 ATTATCTATGAAGAAACAACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTTT 1327
QY 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
Db 1328 CCCCAAGCTCAGAGAGGAAATTTGAGTTTATTTTTTGGCTCTGAAATGCCAAAACAGGTTTC 1387
QY 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
Db 1388 CGTCATTTATACAAAATTTACATCTTATTTTAAAGGAAGCAAAATATAAAGCATCCAGTGGT 1447
QY 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 500
Db 1448 GGGCTGCTGCTCCAAAGTGATTTCAAGTGTCTTATCAAGAGGAGATAGCAATTTACCAGT 1507
QY 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValAlaArg 520
Db 1508 GGTGAATGGGAAGTTCCTGGCCGCGATGATCTAATATCCAACTTGATGAAGTCAGAAG 1567
QY 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisIleLeuTyrValValSer 540
Db 1568 CTGGTATATTTTGAAGGCCAAAGATCTCCCTTTAGAGCATCCTGTACGTAGTCACT 1627
QY 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
Db 1628 TAGCTAAATCTCGGAGAGGTGACAAAGCTGACAGCCGTGGCTACATCACATCTTGCTGC 1687
QY 561 IleSerGlnHisCysAspPheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
Db 1688 ATCAGTCAGCACTGTGACTCTTTATTAAGTAAGTATAGTAACCAAGAAATCCACACTGT 1747
QY 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
Db 1748 GTGTCCCTTTACAAGCTATCAAGTCTCTGAAGATGACCCAACTTTGCAAAAACAAAGGAATTT 1807
QY 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 620
Db 1808 TGGGCCACCATTTTGGATTACAGCAGGTCTCTTCTCTGACTATACTCTCCAGAAAATTTTC 1867
QY 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
Db 1868 TCTTTTGAAGTACTACTGGATTTTACATTTGATGGGATGCTCTACAAGCCTCATGATCTA 1927
QY 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyProGlnValGlnLeu 660
Db 1928 CAGCTCGAAGAAATATCTCTGCTGTTCTATATATGTTGGTCTCTCAGGTGAGTTG 1987
QY 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680
Db 1988 GTGAATAATTCGAATTTAAAGGAGTCAAGTATTTCCGTTGAATACCCCTAGCTCTCAGGT 2047
QY 681 TyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 700
Db 2048 TATGTGGTTGTAGTATAGAACACAGGGGATCTCTGTCCAGGAGGCTTAAATTTGAAGGC 2107
QY 701 AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr 720
Db 2108 GCCTTTAAATATAAATGGGTCAAAATAGAAATAGACATCAGGTGGAAGACTCCCAATAT 2167
QY 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyr 740

Db 2168 CTAGCTTCTCGATATGATTTTCATTGACTTAGATCGTGTGGGCATCCACGGCTGCTCTAT 2227
Qy 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
Db 2228 GGAGGATACCTCTCCCTGATGGCAATTAATGCGAGGTCTAGATATCTTCAGGGTGGCTATT 2287
Qy 761 AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780
Db 2288 GCTGGGGCCCCAGTCACTCTGTGTGATCTTCTATGATACAGGATACACCGAAGCTTATATG 2347
Qy 781 GlyHisProAspGlnAsnGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
Db 2348 GGTCACTCTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAA 2407
Qy 801 LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
Db 2408 AAGTTCCTCTCGAACCAATCGTTTACTGCTCTTACATGGTTTCTGTGATGAGATGTC 2467
Qy 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
Db 2468 CATTTTGCACATACCACTATATTACTGAGTTTTTTTAGTGAGGGCTGGAAAGCCATATGAT 2527
Qy 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
Db 2528 TTACAGATCTATCTCTCAGGAGACACAGCATAGAGTTCTCTGAATCGGGAGAACATTAT 2587
Qy 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLys 880
Db 2588 GAACCTGCATCTTTTGGCACTACCTTCAAGAAAACTTGGATCAGATTTCTGCTCTATAAA 2647
Qy 881 ValIle 882
Db 2648 GTGATA 2653

RESULT 2

US-10-070-464-2
; Sequence 2, Application US/10070464
; Patent No.: 6981564
; GENERAL INFORMATION:
; APPLICANT: ABBOTT, Catherine' Anne
; APPLICANT: CORRELL, Mark Douglas
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES
; FILE REFERENCE: GH-007
; CURRENT APPLICATION NUMBER: US/10/070,464
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: PCT/AU00/01085
; PRIOR FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: AU PQ5709
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: AU PQ2762
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 3120
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-070-464-2

Alignment Scores:
Pred. No.: 0 Length: 3120
Score: 4700.00 Matches: 882
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 3 Gaps: 0

US-10-825-632-1 (1-882) x US-10-070-464-2 (1-3120)

Qy 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
Db 214 ATGGCAGCAGCAATGGAAACAGACAGCTGGGTGTGAGATATTGAAACTGGCGACTGT 273

Qy 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
Db 274 GAGGAGAAATATTGAATCACAGGATCGGCCTAAATTTGAGCCCTTTTATGTTGAGCGGTAT 333
Qy 41 SerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
Db 334 TCTTGGAGTCACTTTAAAAGCTGCTTGCAGATACAGAAAAATATCATGGCTACATGATG 393
Qy 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
Db 394 GCTAAGGACCCACATGATTTTCAATGTTGTGAAGAGGAATGATCCAGATGGACCTCAATCA 453
Qy 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
Db 454 GACAGAATCTATTACCTTTCCTCATGCTCTGGTGAGAACAGAGAAATACATGTTTTATTCT 513
Qy 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
Db 514 GAAATTTCCAAAACATATCAATAGACAGCAGTCTTAATGCTCTCTTGGGAAGCCTCTTTTG 573
Qy 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg 140
Db 574 GATCTTTTTCAGGCAACACATGGAATATGTAATGTTCTTCGAGAAGAAGAACTATTAAAGA 633
Qy 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
Db 634 GAAAGAAAACCATTTGGGAACAGTCGGAATTCCTTTACGATTATCACCAAGAGAAAGTGA 693
Qy 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
Db 694 ACATTTCTGTTTCAAGCCGGTAGTGGAAATTTATCACGTAAAGATGAGAGGGCCACAAGGA 753
Qy 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
Db 754 TTTACGCAACACCTTTTAAGGCCCAATCTAGTGGAAACTAGTTGTCTCCACACATACCGATG 813
Qy 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
Db 814 GATCCAAAATTTATGCCCGCTGATCCAGACTGGGATGCTTTTATACATATGACACAGATATT 873
Qy 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu 240
Db 874 TGGATATCTAACATCTGAACCCAGAGAGAAGAGAGACTCACTTATGTGCACAAATGAGCTA 933
Qy 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu 260
Db 934 GCCAACATGGAAGAAGATGCCAGATCAGCTGAGTGGCTACCTTTGTCTCCAGAAAGAA 993
Qy 261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280
Db 994 TTTGATAGATATTCTGGCTATTGGGTGCTCCAAAAGCTGAAACAACTCCAGTGGTGGT 1053
Qy 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHisVal 300
Db 1054 AAAATTTCTTAGAATTTCTATGCAAGAAATGATGATCTGAGGTGGAAATTTATTCATGTT 1113
Qy 301 ThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThr 320
Db 1114 ACATCCCTTATGTTGGAAACAAGGAGGCGAGATTCATTCCTGTTATCTCTAAACAGGTACA 1173
Qy 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
Db 1174 GCAAACTCTAAAGTCACTTTTAAAGATGTGAGAAATTAATGATGTGATGCTGAAGAGGAGATC 1233
Qy 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
Db 1234 ATAGAGTTCATAGTAAGGAACATAATTCACCTTTTGAGATCTCTATTGAGAGGATTGAA 1293
Qy 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
Db 1294 TATATTGCCAGACTGGATGGACTCTGAGGGGAAATATGCTTGGTCCATCTACTAGAT 1353
Qy 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400

Db 334 TCCTGGAGTCAGCTTAAAGAGCTGCTTCCGATACACAGAAATATCATGCTACATGATG 393
Qy 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
Db 394 GCTAAGGCCACCATGATTTTCATGTTGTGAAGAGGAATGATCCAGATGACCTCATTTCA 453
Qy 81 AspArgLysTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
Db 454 GACAGAATCTATTACCTTGGCCATGCTGCTGGTGAGAACAGAGAAATACATGTTTATTCT 513
Qy 101 GluLysProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTyrLysProLeuLeu 120
Db 514 GAATTCCTCCAAACATCAATAGAGCAGCAGTCTTAATGCTCTCTTGGAGCCTCTTTTG 573
Qy 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140
Db 574 GATCTTTTTCAGGCACACTGGACTATGGAATGATTTCTCGAGAAGAAGAACTATTAAAG 633
Qy 141 GluArgLysArgLysGlyThrValGlyLysLeuSerTyrAspTyrHisGlnGlySerGly 160
Db 634 GAAGAAGAACCGATTTGGAAACAGTCGGAATTCCTTACGATTTATCACCAGGAAGTGA 693
Qy 161 ThrPheLeuPheGlnAlaGlySerGlyLysThrHisValLysAspGlyGlyProGlnGly 180
Db 694 ACATTTCTGTTTCAAGCCGCTAGTGGAAATTTATCAGTAAAGATGGAGGGCCACAGGA 753
Qy 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnLeuArgMet 200
Db 754 TTTACGCAACAACTTTAAGCCCAATCTAGTGGAACTAGTTGTGCCAACATACGGATG 813
Qy 201 AspProLysLeuCysProAlaAspProAspTyrIleAlaPheIleHisSerAsnAspIle 220
Db 814 GATCCAAATATTATGCTGCTGCTGATCCAGACTGGATTCCTTTATATACATAGCAACCATATT 873
Qy 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu 240
Db 874 TGGATATTCTAATCATGCTNACCAGAGAAGAAAGAGAGCTCATTTATGTGCACATGAGCTA 933
Qy 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu 260
Db 934 GCCAATGGAAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTGCTCCCAAGAGAA 993
Qy 261 PheAspArgTyrSerGlyTyrTrpTyrCysProLysAlaGluThrThrProSerGlyGly 280
Db 994 TTTGATAGATATTCTGGCTATTGGTGGTGTCCAAAGAGCTGAACAACTCCAGTGGTGT 1053
Qy 281 LysIleLeuArgLysLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300
Db 1054 AAAATCTTAGAATTTCTATATGAAGAAATGATGAATCTGAGGTGGAATATTATCATGTT 1113
Qy 301 ThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThr 320
Db 1114 ACATCCCTATGTTGGAACCAAGAGGGCAGATTCATTCGTTATCTTAAACAGGTACA 1173
Qy 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
Db 1174 GCAATCTCTAAAGTCACCTTTTAAGATGTCAGAAATAATGATGATGCTCCAGGAAGGATC 1233
Qy 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
Db 1234 ATAGATGTCATAGATAAGGAATAATTCACCTTTTGGAGATTCATTTGAAGGAGTTGAA 1293
Qy 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
Db 1294 TATATTGCAGAGCTGGATGAGCTCTCTGAGGGAATAATGCTTGGTCCATCTCTACTAGAT 1353
Qy 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
Db 1354 CGCTCCAGACTCGCTTACAGATAGTGTGTGATCTCACCTGAAATTTATTTATCCAGTAGAA 1413
Qy 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
Db 1414 GATGATGTTATGGAAGGACAGAGACTCAATTGAGTCAGTGCCTGATTTCTGTGACGCCACTA 1473

Qy 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
Db 1474 ATTATCTATGAAGAACAACACACATCTGGATAAATATCCATGACATCTTTTCATGTTTTT 1533
Qy 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
Db 1534 CCCCAAGTCCAGAGAGAAATGAGTTATTTTTTTCCTCTGAATGCCAACAACAGGTTTC 1593
Qy 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
Db 1594 CGTCATTTATACAAAATTTACATCTATTTTAAAGGAAAGCAAAATATAAACGATCCAGTGT 1653
Qy 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 500
Db 1654 GGGCTCGCTGCTCCCAAGTGTCTTCAAGTGTCTTATCAAGAGGAGATAGCAATTTACCACT 1713
Qy 501 GlyGluTrpGluValLeuGlyVarHisGlySerAsnIleGlnValAspGluValArgArg 520
Db 1714 GGTGAATGGGAAGTTCTTGGCCGCGATGGATCTAATATCCAGTTGATGAGTCAGNAGG 1773
Qy 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer 540
Db 1774 CTGGTATATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTACGTAGTCACT 1833
Qy 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
Db 1834 TAGTAAATCTCTGGAGAGGTGCAAGGCTGACTGACCGTGGCTACTCACATTTCTTGCTGC 1893
Qy 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
Db 1894 ATCAGTCAGCATCTGTGACTTCTTTATAGTAGATATAGTAACCAAGAAATCCACACTGT 1953
Qy 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
Db 1954 GTGTCCCTTTACAAGCTATCAAGTCTCTGAAGATGCCCAACTTGCACAAACAAAGGAATTT 2013
Qy 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePhe 620
Db 2014 TGGGCCACCATTTTGGATTCAGCAGTCTCTCTGCTGACTATATCTCTCCAGAAATTTTC 2073
Qy 621 SerPheGluSerThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
Db 2074 TCTTTTGAAGTACTACTGGAATTTACATTTGATGGATGCTCTCAAGCCCTCATGATCTA 2133
Qy 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGly--ProGlnValGlnL 660
Db 2134 CAGCCTGGAAAGAAATATCTCTACTGCTGTTCTATATATGTTGTTCTCTCAGGTGCAGT 2193
Qy 660 euValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuG 680
Db 2194 TGGTGAATTAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTTAGCCTCTCTAG 2253
Qy 680 LysTyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluG 700
Db 2254 GTTATGTGGTGTGTAGTGATAGACACAGGGGATCCTGTCCACCGAGGCTTTAAATTTGAAG 2313
Qy 700 LysAlaPheLysTyrLysMetGlyGlnIleGluLeuAspAspGlnValGluGlyLeuGln 720
Db 2314 GGGCTTTTAAATATAAATGGTCAATAGAAATTTGACGATCAGGTGGAAGGACTCCAAT 2373
Qy 720 TyrLeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSer 740
Db 2374 ATCTAGTCTTCGATATCATTTTCTTACTTAGATCTGTGGGCATCCACGGCTGTCTCT 2433
Qy 740 TyrGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaI 760
Db 2434 ATGGAGGATACCTCTCCCTGATGCAATTAATGCAAGGTCAGATATCTTCAGGGTTGCTA 2493
Qy 760 LeuAlaGlyValProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrM 780
Db 2494 TTGCTGGGGCCCCCAGTCTCTCTGTGGATCTTCTATGATACAGGATACACGGAACGTTATA 2553

Qy	780	etGlyHisProAspGlnAenGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaG	800
Db	2554	TGGGTCAACCTGACCAAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAG	2613
Qy	800	lulysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAenV	820
Db	2614	AAAAGTTCCCTCTGAACCAAAATCGTTTACTGCTCTTACATGGTTTCCGTGAGAAATG	2673
Qy	820	alHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrA	840
Db	2674	TCCATTTGGACATACCAAGTATATTACTGAGTTTITTAGTGAGGCTGGAAGCCATATG	2733
Qy	840	spLeuGlnIleTyrProGlnGluArgHisIleArgValProGluSerGlyGluHisT	860
Db	2734	ATTTACAGATCTATCTCTCAGAGAGACACAGATAGAGTTCTTGAAATCGSGAGAACATT	2793
Qy	860	YrGluLeuHisLeuLeuHisTyrIleuGlnGluAsnLeuGlySerArgIleAlaLeuL	880
Db	2794	ATGAACATGCATCTTTTGGCACTACCTTCAAGAAACCTTGGATCAGTATTGCTCTTAA	2853
Qy	880	YsValIle 882	
Db	2854	AAGTGATA 2861	
RESULT 4			
US-09-976-594-1103			
; Sequence 1103, Application US/09976594			
; Patent No. 6673549			
; GENERAL INFORMATION:			
; APPLICANT: Furness, Michael			
; APPLICANT: Buchbinder, Jenny			
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS			
; FILE REFERENCE: PA-0041 US			
; CURRENT APPLICATION NUMBER: US/09/976,594			
; CURRENT FILING DATE: 2001-10-12			
; PRIOR APPLICATION NUMBER: 60/240,409			
; PRIOR FILING DATE: 2000-10-12			
; NUMBER OF SEQ ID NOS: 1143			
; SOFTWARE: PERL Program			
; SEQ ID NO 1103			
; LENGTH: 2797			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
; FEATURE:			
; NAME/KEY: misc feature			
; OTHER INFORMATION: Incyte ID No. 6673549 977951.1			
US-09-976-594-1103			
Alignment Scores:			
Pred. No.:	0	Length:	2797
Score:	4443.50	Matches:	840
Percent Similarity:	99.5%	Conservative:	0
Best Local Similarity:	99.5%	Mismatches:	1
Query Match:	94.5%	Indels:	3
DB:	3	Gaps:	1
US-10-825-632-1 (1-882) x US-09-976-594-1103 (1-2797)			
Qy	42	TpSerGlnLeuLysLeuLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMetAla	61
Db	3	TGAGTCACTTAAAAGCTGCTTCCGATACCAAGAAATATCATGCTACATGATGGCT	62
Qy	62	LysAlaProHisAspPheMetPheValLysArgHisAspProAspGlyProHisSerAsp	81
Db	63	AAGGCACCATGATTTTCAATTTTGTGAAGAGATGATCCAGATGGACCTCATTCAGAC	122
Qy	82	ArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSerGlu	101
Db	123	AGAAATCTATTACTTGGCCATGCTCGTGAGAACACAGAAATACACATGTTTATTCTGAA	182
Qy	102	IleProLysThrIleAsnArgAlaValLeuMetLeuSerTyrLysProLeuLeuAsp	121
Db	183	ATTCCTCAAAATCATCAATAGACGACGCTCTTAATGCTCTCTTGGAGGCTCTTTGGAT	242

Qy 480 lYcLYLeuProAlaProSerAspPheLysCysProLysGluGluLeuAlaLeuThrS 500
Db 1323 GTGGGCTGCTGCTCAAGTGATTTCAAGTCTCTATCAAGAGGAGATAGCAATPACCA 1382
Qy 500 erGLyLutrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgA 520
Db 1383 GTGGTGAATGGGAAGTTCTTGGCCGGCATGGATCTAATATCCAAAGTTGATGAAGTCAGAA 1442
Qy 520 rGLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValS 540
Db 1443 GGCTGCTATATTTGAAGGACCAAGACTCCCTTTAGAGCATCCTCTGCTAGTGTCA 1502
Qy 540 erTyrValLeuProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysC 560
Db 1503 GTTACGTAATCTCGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACATTCCTTGT 1562
Qy 560 ystleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisC 580
Db 1563 GCATCAGTCAGCACTGTGACTCTTTTATAAGTAAGTATAGTAACCAAGAGATCCACACT 1622
Qy 580 ySValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluP 600
Db 1623 GTGTGTCCTTTACAGCTATCACTCTGAAGATGACCCCACTTGCACAAACAAGAAAT 1682
Qy 600 heTPrAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIleP 620
Db 1683 TTTGGGCCACCAATTTGGATTGAGCAGGTCTCTTCTGCTACTACTCTCTCCAGAAATTT 1742
Qy 620 heSerPheGluSerThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspL 640
Db 1743 TCTCTTTTGAAGTAGTACTACTGGATTACATTTGATATGGGATGCTCTCAAGCCCTCATGATC 1802
Qy 640 euGlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnL 660
Db 1803 TAGCCTCGAAAGAAATATCTACTGCTGCTTCATATATATGCTGCTCAGTGCAGT 1862
Qy 660 euValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuG 680
Db 1863 TGGTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTCAATACCTAGCCTCTCTAG 1922
Qy 680 lYTrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluG 700
Db 1923 GTTATGTGTTGTAGTGATAGACACAGGGGATCCTGTCAACCGAGGGCTTTAAATTTGAAG 1982
Qy 700 lYAlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGlnGlyLeuGlnT 720
Db 1983 GCSCCTTTAAATATAAATGGGTCAAAATAGAAATGACATGAGTGGGAAGGACTCAAT 2042
Qy 720 YrLeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerT 740
Db 2043 ATCTAGCTTCTCGATATGATTTTCATTTAGCTTAGATCGTGGGCATCCAGGCTGCTCT 2102
Qy 740 YrGlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaI 760
Db 2103 ATGGAGGATACCTCTCCCTGATGGCAATTAATGCAGAGTCCAGATATCTTCAGGGTGTCTA 2162
Qy 760 leAlaGlyAlaProValThrLeuTyrIlePheTyrAspThrGlyTyrThrGluArgTyrW 780
Db 2163 TTGCTGGGGCCCCAGTCACTCTGTGGATCTCTATGATACAGGATACACCGAAGCTTATA 2222
Qy 780 etGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaG 800
Db 2223 TGGGTACCTCTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATCGAAGCAG 2282
Qy 800 lulySPhProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnV 820
Db 2283 AAAAGTTCCCTCTCTGAACCAATCGTTTACTGCTTACTATGTTCTTCTGGATGAGAAATG 2342
Qy 820 alHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrA 840
Db 2343 TCCATTTTGCACATACCATATATTTACTGAGTTTTTTTAGTGAGGGCTGGAAGCCCATATG 2402
Qy 840 spLeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHist 860

Db 2403 ATTTACAGATCTATCTCTCAGAGAGACACACCATPAAGTTCCTGAATCGGAGACATT 2462
Qy 860 YrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeuL 880
Db 2463 ATGAACCTGCATCTTTTGACCTACCTTCAAGAAAAACCTTGGATCAGCTATTGCTGCTCTAA 2522
Qy 880 ySValIle 882
Db 2523 AAGTGATA 2530
RESULT 5
US-09-976-674-22
; Sequence 22, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCES: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 22
; LENGTH: 4685
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-976-674-22
Alignment Scores:
Pred. No.: 0 Length: 4685
Score: 4385.50 Matches: 834
Percent Similarity: 94.6% Conservative: 0
Best Local Similarity: 94.6% Mismatches: 1
Query Match: 93.3% Indels: 48
Gaps: 1
US-10-825-632-1 (1-882) x US-09-976-674-22 (1-4685)
Qy 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
Db 214 ATGCAGCAGCAATGGAAACAGACAGCTGGGTGTGAGATATTTGAACTCGGAGCTGT 273
Qy 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
Db 274 GAGGAGAAATATTGAATCACAGGATCGGCTAAATTTGGAGCCCTTTTATGTTGAGCGGTAT 333
Qy 41 SerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
Db 334 TCTGTGAGTCACTTAAAGGCTGCTGCCATATCCAGAAATAATCATGCTGATGATG 393
Qy 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
Db 394 GCTAAGCACCACATGATTTTCATGTTGTGAAGGAGATGATCCAGATGACCTCATTC 453
Qy 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
Db 454 GACAGAAATCTATTACCTTGCCATGCTGCTGAGAACACAGAGAAATAACACTGTTTATCT 513
Qy 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
Db 514 GAAATTTCCCAAAATATCAATAGACAGCAGCTCTTAATGCTCTCTTGGAGCCCTCTTTTG 573
Qy 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg 140
Db 574 GATCTTTTTCAGGCAACACTGAGCTATGGAATGATTTCTCGAAGAGAGAACTATTAGA 633
Qy 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160


```
QY 541 TyrValIleProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
DB 1834 TACGTAAATCTCGAGAGGTGACAGGCTGACTGACCGTGGCTACTACATCTTCGCTGC 1893
QY 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
DB 1894 ATCAGTCAGCACTGTGACTCTCTTATAAGTAGTAGTAAACCAAGAAATCCACACTGT 1953
QY 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
DB 1954 GTGTCCCTTTACAAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAAAACAAAGAAATTT 2013
QY 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 620
DB 2014 TGGGCCACCATTTTGGATTTCAGCAGGTCTCTCTCTGACTATACTCTCCAGAAATTTTC 2073
QY 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
DB 2074 TCTTTTGAAGTACTACTGGATTTCATTGTATGGATGCTCTCAAGCCCTCATGATCTA 2133
QY 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 660
DB 2134 CAGCTTGGAAAGAAATATCTACTGTCTCTCATATATGTTGGTGC----- 2180
QY 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680
DB 2180 ----- 2180
QY 681 TyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 700
DB 2180 ----- 2180
QY 701 AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr 720
DB 2181 -----GGTCAATAGAAATTTGACGATCAGGTGGAGGACTCCAAATAT 2222
QY 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyr 740
DB 2223 CTAGCTTCTCGATATGATTTTCATTGACTTAGATCGTGTGGGCATCCACGGCTGGTCCCTAT 2282
QY 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
DB 2283 CGAGGATACCTCTCCCTGATGGCAATTAAATGCAGAGGTTCAGATATCTTCAGGGTTCGTATT 2342
QY 761 AlaGlyAlaProValThrLeuTyrIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780
DB 2343 GCTGGGGCCCGACCTCTCTGGATCTTCTATGATACAGGATACACGGAAAGTTATATG 2402
QY 781 GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
DB 2403 GGTCAACCTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAA 2462
QY 801 LysPheProSerGluProAsnArgLeuLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
DB 2463 AAGTTCCTCTGAAACCAATCGTTTACTGCTCTTACATGGTTTCTGGATGAGAAATGTC 2522
QY 821 HisPheAlaHisThrSerIleLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
DB 2523 CATTTTGCATACACGATATATTAAGTTTTTACTGAGGGCTGGAAAGCCATATGAT 2582
QY 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
DB 2583 TTACAGATCTATCTCTCAGAGAGACAGCATAGATTCTCGAATCGGAGAACATATAT 2642
QY 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeuLys 880
DB 2643 GAACTGCAATCTTTTGCATCTCTTCAAGAAACCTTGGATACAGTATTCGCTCTCAAAA 2702
QY 881 ValIle 882
DB 2703 GTGATA 2708
```

RESULT 7

```
US-09-976-674-8
; Sequence 8, Application US/09976674
; Patent No. 684190
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 4523
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-976-674-8

Alignment Scores:
Pred. No.: 0 Length: 4523
Score: 0 902.50 Matches: 780
Percent Similarity: 88.4% Conservative: 0
Best Local Similarity: 88.4% Mismatches: 1
Query Match: 87.1% Indels: 102
DB: 3 Gaps: 1

US-10-825-632-1 (1-882) x US-09-976-674-8 (1-4523)
QY 1 MetAlaAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
DB 214 ATGGCAGCAGCATATGGAACACAGACAGCTGGTGTTCAGATATTTGAACTCGGACTGT 273
QY 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
DB 274 GAGGAGAAATATTGAATACAGATCGGCTAAATGGAGCTTTTATATGTTGACCGGTAT 333
QY 41 SerTyrSerGlnLeuLysLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
DB 334 TCTCGAGTCAGCTTAAAGAGCTGCTTGCCTATACAGAAATATCATGCTACATGATG 393
QY 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
DB 394 GCTAAGGCACCACTATGATTTTCATGTTGTGAAGAGGAATGATCCAGATGGACCTCATTC 453
QY 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
DB 454 GACAGAAATCTATTACCTTGCCTATGCTGTGGTGAGAACAGAGAAATACACATGTTTATTC 513
QY 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTyrLysProLeuLeu 120
DB 514 GAAATTTCCCAAACTATCAATAGACAGCAGCTCTTAATGCTCTCTTGGAAAGCCCTCTTTTG 573
QY 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg 140
DB 574 GATCTTTTTCAGCAACACTGGACTATGGAATGTTATTTCTCGAGAAGAAAGAACTATTAAAG 633
QY 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
DB 634 GAAAGAAACCACTATGGAACAGTTCGAATTCCTCTTACGATTATACCAAGAAAGTGA 693
QY 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
DB 694 ACATTTCTGTTTCAGCCGCTAGTGAATTTATCCGTAAAGATGAGAGGCCACCAAGGA 753
QY 181 PheThrGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
DB 754 TTTTACCAACCAACCTTTTAAAGCCCACTAGTGGAAACTAGTTCTCCCAACATACGATG 813
QY 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
```


APPLICANT: Riviere, Pierre
APPLICANT: Junien, Jean-Louis
TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
FILE REFERENCE: 70669
CURRENT APPLICATION NUMBER: US/09/976,674
CURRENT FILING DATE: 2001-10-12
PRIOR APPLICATION NUMBER: US 60/240,117
PRIOR FILING DATE: 2000-10-12
NUMBER OF SEQ ID NOS: 61
SOFTWARE: PatentIn version 3.1
SEQ ID NO 14
LENGTH: 4309
TYPE: DNA
ORGANISM: Homo sapiens
US-09-976-674-14

Alignment Scores:
Pred. No.: 0 Length: 4309
Score: 3661.50 Matches: 708
Percent Similarity: 80.1% Conservative: 9
Best Local Similarity: 80.1% Mismatches: 1
Query Match: 77.9% Indels: 176
DB: 3 Gaps: 1

US-10-825-632-1 (1-882) x US-09-976-674-14 (1-4309)

QY	1	MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys	20
DB	214	ATGGCAGCAGCAATGGAAACAGACAGCTGGGTGGATATTTGAGATATTTGAAACTGGCGACTGT	273
QY	21	GluGluAenIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr	40
DB	274	GAGGAGAATATTGAATCACAGGATCGGCCTAAATTTGGAGCCTTTTATGTGGCGGTAT	333
QY	41	SerTrpSerGlnLeuLysLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet	60
DB	334	TCTGGAGTCAGCTTAAAGAGCTGCTTGGCGATACACAGAAATATCATGGCTACATGATG	393
QY	61	AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer	80
DB	394	GCTAAGGCACCATGATTTTCATGTTGTGAGAGGAAATGATCCAGATGGACCTCATTTCA	453
QY	81	AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer	100
DB	454	GACAGAATCTATTACTTGGCCATGTCGTGGAGACGAGTCTTAATGCTCTTTGGAGGCTCTTTTG	513
QY	101	GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu	120
DB	514	GAAATTTCCAAACTATCATAGAGCAGCAGTCTTAATGCTCTTTGGAGGCTCTTTTG	573
QY	121	AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg	140
DB	574	GATCTTTTTCAGGCACACTGGACTATGGAAATGATTTCTCGAAGAAGAACTATTAGA	633
QY	141	GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly	160
DB	634	GAAAGAAAACGGATTTGGAACAGTCGGAATTTATCAGCTTAAAGATGGAGGCGGCACAAAGA	693
QY	161	ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly	180
DB	694	ACATTTCTGTTTCAAGCGGTAGTGGAAATTTATCAGCTTAAAGATGGAGGCGGCACAAAGA	753
QY	181	PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet	200
DB	754	TTTACGCACACACCTTTAAGGCCCAATCTAGTGGAACTAGTTGTCCCAACATACGATG	813
QY	201	AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle	220
DB	814	GATCCAAAATTTATGCCCTGCTGATCCAGACTGGATTTGCTTTTATACATAGCAACGATATT	873
QY	221	TrpIleSerAsnIleValThrArgGluGluArgGluLeuThrTyrValHisAsnGluLeu	240
DB	874	TGGATATCTACATGTTATACCAAGAGAGAGAGACTCACTTATGTGCACATGAGCTA	933

QY	241	AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu	260
DB	934	GCCAAATCGAAGAAAGATGCCAGATCAGCTGGAGTCGTACCTTTGTTTCCNAGAAGAA	993
QY	261	PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly	280
DB	994	TTTGATAGATATTCTGGCTATTGGTGTGCTCCAAAGCTGAAACCACTCCCACTGGTGGT	1053
QY	281	LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal	300
DB	1054	AAAAATCTTGAATTTCTATATGAAGAAATGATGAATCTGAGGTGGAATATTATTCTGTT	1113
QY	301	ThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThr	320
DB	1114	ACATCCCTATGTTGGAACAAGAGGCGAGATTCTCCGTATCTCTTAAACAGGTACA	1173
QY	321	AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle	340
DB	1174	GCAAAATCCTAAAGTCACTTTTAAAGATGTCAGAAATAATGATTGATGCTGAAGGAAGATC	1232
QY	341	IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu	360
DB	1232	-----	1232
QY	361	TyrIleAlaArgAlaGlyTyrThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp	380
DB	1232	-----	1232
QY	381	ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu	400
DB	1232	-----	1232
QY	401	AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu	420
DB	1232	-----	1232
QY	421	IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe	440
DB	1232	-----	1232
QY	441	ProGlnSerHisGluGluGluIlePheAlaSerGluCysLysThrGlyPhe	460
DB	1232	-----	1232
QY	461	ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly	480
DB	1232	-----	1232
QY	481	GlyLeuProAlaProSerAspPheLysCysProIleLysGluIleAlaIleThrSer	500
DB	1232	-----	1232
QY	501	GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg	520
DB	1233	-----CAAGTTGATGAAGTCAGAAAGG	1253
QY	521	LeuValTyrPheGluGlyThrIleAspSerProLeuGluHisHisLeuTyrValValSer	540
DB	1254	CTGTATATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCCTGTACGTAGTCAGT	1313
QY	541	TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys	560
DB	1314	TACGTAAATCCTGGAGAGGTGACAAAGCTGACTGACCTGGCTACTCACATTCTTCTGTC	1373
QY	561	IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys	580
DB	1374	ATCAGTCAGACTGTGACTTCTTATAAGTAAGTATAGTACCAAGAAATCCACTGT	1433
QY	581	ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe	600
DB	1434	GTGTCCCTTTTACAAGCTATCAAGTCTCAAGATGACCCACTTGCATAAACAAGGAATTT	1493

Qy 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 620
Db 1494 TGGGCCACCAATTTTGATTCAGCAGGCTCTCTCTGACTATATCTCTCCAGAAATTTTC 1553
Qy 621 SerPheGluSerThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
Db 1554 TCITTTGAAGTACTACTGATTTACATTTGATGGATGCTCTCAAGCCCTCATGATCTA 1613
Qy 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGly--ProGlnValGlnL 660
Db 1614 CAGCCTGGAAGAAATATCTACTGTCTGCTTCATATATGCTGCTCTCTCAGTGCACT 1673
Qy 660 euValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuG 680
Db 1674 TGGTGAATATCGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTAGCTCTCTAG 1733
Qy 680 lYrYrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluG 700
Db 1734 GTTATGTTGTTGATGTAGACACAGGGGATCTGTCCAGGGGCTTAAATTTGAAG 1793
Qy 700 lYAlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnT 720
Db 1794 GCGCTTTAATATAAATGGTCAATAGAAATTCAGCATCAGTGGAGGACTCCAAT 1853
Qy 720 YrLeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerT 740
Db 1854 ATCTAGCTTCTCGATATGATTTCAITGATCTAGATCGTGTGGGCATCCAGGCTGGTCT 1913
Qy 740 YrGlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaI 760
Db 1914 ATGGAGGATACCTTCCCTCGATGGCAATTAATGACAGAGTTCAGATATCTTCAGGGTGGCTA 1973
Qy 760 leAlaGlyAlaProValThrLeuTyrIlePheTyrAspThrGlyTyrThrGluArgTyrM 780
Db 1974 TTGCTGGGGCCCCGACGTCATCTGTGAATCTTCTATGATACAGGATACACGGACGTTATA 2033
Qy 780 etGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaG 800
Db 2034 TGGGTCACTCTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAG 2093
Qy 800 luyysPheProSerGluProAsnArgLeuLeuLeuLeuHisGlyPheLeuAspGluAsnV 820
Db 2094 AAAAGTTCCCTCTGAACCAATCGTTTACTCTCTTACATGGTTCCTCGATGAGAATG 2153
Qy 820 alHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrA 840
Db 2154 TCCATTTTGACATACAGTATATTAATGAGTTTTTTAGTGGGGCTGGAAAGCCATATG 2213
Qy 840 spLeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHist 860
Db 2214 ATTTACAGATCTATCTCTCAGGAGACACAGCATTAAGAGTTCTCGAATCGGAGAACATT 2273
Qy 860 YrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuL 880
Db 2274 ATGAACATGCACTCTTTTGCACTACTTCAAGAAACCTTGGATCAGTATGCTGCTCTAA 2333
Qy 880 ySValIle 882
Db 2334 AAGTGATA 2341

RESULT 9

US-09-976-674-4
; Sequence 4, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE OF INVENTION: 70669
; CURRENT APPLICATION NUMBER: US/09/976, 674
; CURRENT FILING DATE: 2001-10-12

; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 4
; LENGTH: 2617
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-976-674-4

Alignment Scores:
Pred. No.: 1-92e-315 Length: 2617
Score: 2870.00 Matches: 517
Percent Similarity: 77.5% Conservative: 134
Best Local Similarity: 61.5% Mismatches: 187
Query Match: 61.1% Indels: 2
DB: 3 Gaps: 2

US-10-825-632-1 (1-982) x US-09-976-674-4 (1-2617)

Qy 35 PheTyrValGluArgTyrSerTrpSerGlnIleuLysLysLeuLeuAlaAspThrArgLys 54
Db 80 TTCAGGTGCAGAAAGCACTCGTGGACGCGGCTCCGGAGCATCATCCAGCGCGCGCAAG 139
Qy 55 TyrHisGlyTyrMetMetAlaLysAlaProHisAspPheMetPheValLysArgAsnAsp 74
Db 140 TACTCGGGCTCTATTGTTCAACAAAGCGGCCCAAGCACTTCCAGTTTGTGCAAGAACCGAT 199
Qy 75 ProAspGlyProHisSerAspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGlu 94
Db 200 GAGTCTGGGGCCCACTCCACCGCTCTACTACCTGGGAATGCCATATGGCAGCGGAGAG 259
Qy 95 AsnThrLeuPheTyrSerGluIleProLysThrIleAsnArgAlaAlaValLeuMetLeu 114
Db 260 AACTCCCTCTCTACTCTGAGATTCCCAAGAGGTCCGGAAGAGGCTCTGCTCTCTG 319
Qy 115 SerTrpLysProLeuLeuAspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArg 134
Db 320 TCTTGGAAAGCAGATGCTGATCATTTCCAGGCGCACCCCACTTCTCTCTCTCTCTCGG 379
Qy 135 GluGluGluLeuLeuArgGluArgLysArgIleGlyThrValGlyIleAlaSerTyrAsp 154
Db 380 GAGAGAGAGCTGCTGAGGAGGCGGAACCCCTGGGGGTCTTCGGCATCACCTCTCTACGAC 439
Qy 155 TyrHisGlnGlySerGlyThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLys 174
Db 440 TTCCACAGCAGAGATGGCTCTCTCTTCCAGGCGCACCAAGCGCTCTTCCACTGTCTCGC 499
Qy 175 AspGlyGlyProGlnGlyPheThrGlnGlnProLeuArgProAsnLeuValGluThrSer 194
Db 500 GACGCGGCAAGAACCGCTTCATGGTGTCTCCCTATGAAACCGCTGGAAATCAAGACCCAG 559
Qy 195 CysProAsnIleArgMetAspProLysLeuCysProAlaAspProAspTrpIleAlaPhe 214
Db 560 TGCTCAGGGCCCCGAGTGGACCCCAAAATCTGCCCTGCGACCCCTTCTCTCTCTCTCTTC 619
Qy 215 IleHisSerAsnAspIleTrpIleSerAsnIleValThrArgGluGluArgArgLeuThr 234
Db 620 ATCAATAACAGCAGCCTGTGGTGGCCAAACATCGAGACAGCGGAGGAGCGCGCTGACC 679
Qy 235 TyrValHisAsnGluLeuAlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThr 254
Db 680 TTCTGCCACCAAGGTTTATCCAATGCTCTGGATGACCCCAAGTCTGGCGGTGTGGCCACC 739
Qy 255 PheValLeuGlnGluPheAspArgTyrSerGlyTyrTrpTyrCysProLysAlaGlu 274
Db 740 TTGCTCATACAGGAAGATTCGACCGCTTCACTGGGTACTGGTGTGCTGCTGCTGCTGCT 799
Qy 275 ThrThrProSerGlyGly--LysIleLeuArgIleLeuTyrGluGluAsnAspGluSer 293
Db 800 TGGGAAGTTACAGAGGCTCAAGCGCTGCGAATCTCTGTATGAGGAAGTGCATGAGTCC 859
Qy 294 GluValGluIleIleHisValThrSerProMetLeuGluThrArgArgAlaAspSerPhe 313

Db 860 GAGGTGAGGTCAATCAGCTCCCTCTCTCGCGCTAGAGAAAGAGACGCACTCGAT 919
Qy 314 ArgTyrProLysThrGlyThrAlaAsnProLysValThrPheLysMetSerGluLeuMet 333
Db 920 CGGTACCCAGGACAGGACCAAGATCCCAAGATTGCTTGAACCTGGCTGAGTCCAG 979
Qy 334 IleAspAlaGluGlyArgIleLeuAspValIleAspLysGluLeuIleGlnProPheGlu 353
Db 980 ACTGACAGCCAGGACAGATCGTCTCGACCCAGGAGAGGAGCTGGTGCACCCCTTCAGC 1039
Qy 354 IleLeuPheGluGlyValGluTyrIleAlaAsnAlaGlyTyrThrProGluGlyLysTyr 373
Db 1040 TCGCTGTTCGCAAGGTGGAGTACATCGCCAGGCGCGGTGACCCGGATGGCAATAC 1099
Qy 374 AlaTrpSerIleLeuLeuAspArgSerGlnThrArgLeuGluIleValLeuIleSerPro 393
Db 1100 GCCTGGCCATGTTCTCGGACCGCCAGAGTGGCTCCAGCTCGTCTCTCTCCCCCG 1159
Qy 394 GluLeuPheIleProValGluAspValMetGluArgGlnArgLeuIleGlnSerVal 413
Db 1160 GCCTGTTCATCCGACACAGAGAAATGAGGACGCGCTAGCCTCTGCGACGAGCTGC 1219
Qy 414 ProAspSerValThrProLeuIleIleTyrGluGluThrThrAspIleTrpIleAsnIle 433
Db 1220 CCCAGGAATGTCAGCGCTATGTGGTACGAGGAGGTCAACCAAGTCTGGATCAATGTT 1279
Qy 434 HisAspIlePheHisValPheProGlnSerHis---GluGluGluIleGluPheIlePhe 452
Db 1280 CATGACATCTTCTATCCCTCCCCCAATCAGAGGAGAGAGAGGAGCTGTGCTTCTCCGC 1339
Qy 453 AlaSerGluCysValThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLysGlu 472
Db 1340 GCCAATGAATGCAGACCGGCTTCTGCCAATTTGTACAAAGTCAACCGCGCTTTTAAATCC 1399
Qy 473 SerLysTyrLysArgSerSerGlyLeuProAlaProSerAspPheLysCysProIle 492
Db 1400 CAGGCTACGATGAGTGAGCCCTTCAGCCCGGGGAGAGATGAATTTAAGTGCCTCAT 1459
Qy 493 LysGluGluIleAlaIleThrSerGlyGluTrpGluValLeuGlyArgHisGlySerAsn 512
Db 1460 AAGGAAGATTTGCTCTGACCGCGGTGATGGAGGTTTGGCGAGGACGCGCTCCAAG 1519
Qy 513 IleGlnValAspGluValArgLeuValTyrPheGluGlyThrLysAspSerProLeu 532
Db 1520 ATCTGGGTCAATGAGGAGACCAAGCTGTGTACTTCAGGGCCACCAAGACACGCGCTG 1579
Qy 533 GluHisLeuTyrValValSerTyrValAsnProGlyGluValThrArgLeuThrAsp 552
Db 1580 GAGCACCACTCTACGTGTGCTATGAGCGCGCGCGAGATCGTACGCTTCACACG 1639
Qy 553 ArgGlyTyrSerHisSerCysIleSerGlnHisCysAspPhePheIleSerLysTyr 572
Db 1640 CCGGGCTTCTCCCATAGCTGTCTCCATGAGCCAGAACTTCGACATGTTGCTGACCACTAC 1699
Qy 573 SerAsnGlnLysAsnProHisCysValSerLeuTyrLysLeuSerSerProGluAspAsp 592
Db 1700 AGCAGCTGAGCACGCGCGCTCGTGCACGCTCTACAAGCTGAGCGCGCGCGACGAC 1759
Qy 593 ProThrCysLysThrLysGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuPro 612
Db 1760 CCCTGACACAGCAGCCGCTTCTGGGCTAGCATGATGAGGAGGAGCAGCTGCCCGCG 1819
Qy 613 AspTyrThrProGluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGly 632
Db 1820 GATTATGTTCTCCAGAGATCTTCCATTTCCACAGCGCTCGGATGCGGCTCTACGGC 1879
Qy 633 MetLeuTyrLysProHisAspLeuGlnProGlyLysValTyrProThrValLeuPheIle 652
Db 1880 ATGATCTACAGGCCCCACGCTGTGAGCCAGGAGAGACACCCACCGCTCTTTGTA 1939
Qy 653 TyrGlyGlyProGlnValGlnLeuValAsnAsnArgPheLysGlyValLysTyrPheArg 672

Db 1940 TATGAGGCCCCAGGTGAGCTGCTGTAATACTCTTCAAAGGCATCAAGTACTTGGC 1999
Qy 673 LeuAsnThrLeuAlaSerLeuGlyTyrValValValValIleAspAsnArgGlySerCys 692
Db 2000 CTCAACACACTGGCTCCCTTGGCTAGCGCGTGGTGTGATTGACGCGAGGGCTCTGT 2059
Qy 693 HisArgGlyLeuLysPheGluGlyAlaPheLysTyrLysMetGlyGlnIleGluLeuAsp 712
Db 2060 CAGCGAGGGCTTCGGTTGGAAGGGGCTTGAATAACCAATGGCGAGGTGGAGATCGAG 2119
Qy 713 AspGlnValGluGlyLeuGlnTyrLeuAlaSerArgTyrAspPheIleAspLeuAspArg 732
Db 2120 GACCAGGTGAGGCGCTGCAGTTTCGTGGCCGAGAGATATGGCTTTCATGACCTGAGCCGA 2179
Qy 733 ValGlyIleHisGlyTyrSerTyrGlyTyrLeuSerLeuMetAlaLeuMetGlnArg 752
Db 2180 GTTGCCATCCATGGCTGCTCTACGGGGGCTTCTCTCGCTCATGGGGCTATCCACAG 2239
Qy 753 SerAspIlePheArgValAlaIleAlaGlyAlaProValThrLeuTrpIlePheTyrAsp 772
Db 2240 CCCAGGTGTTCAAGGTGGCCATCGCGGTGCCCCGCTCACCGTCTGGATGGCTACGAC 2299
Qy 773 ThrGlyTyrThrGluArgTyrMetGlyHisProAspGlnAsnGluGlnGlyTyrLeu 792
Db 2300 ACAGGTACACTGAGCGCTACATGGAGCTCCCTGAGAACCAACACGACGCTATGAGCG 2359
Qy 793 GlySerValAlaMetGlnAlaGluLysPheProSerGluProAsnArgLeuLeuLeu 812
Db 2360 GTTCCGTGGCCCTGACGTGGAGAGCTGCCCAATGAGCCCAACGCTTGTCTTATCCTC 2419
Qy 813 HisGlyPheLeuAspGluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeu 832
Db 2420 CACGCTCTCTGGAGAGAAAGCTGCACTTTTCCACACAACTTCTCGTCTCCCACTG 2479
Qy 833 ValArgAlaGlyLysProTyrAspLeuGlnIleTyrProGluGluArgHisSerIleArg 852
Db 2480 ATCCGAGCAGGAGAACTTACCAGCTCCAGATCTACCCCAACGAGAGACACAGTATTCG 2539
Qy 853 ValProGluSerGlyGluHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeu 872
Db 2540 TGCCCCGAGTGGCGGAGCACTGTAAGTCACTGCTGTCACCTTTCTACAGGATACCTC 2599

RESULT 10
US-09-976-674-28
; Sequence 28, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 28
; LENGTH: 4219
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-976-674-28

Alignment Scores: 4,33e-315 Length: 4219
Pred. No.: 2870.00 Matches: 517
Score: 77.5% Conservative: 134
Percent Similarity: 61.5% Mismatches: 187
Best Local Similarity: 61.1% Indels: 2
Query Match: 3 Gaps: 2
DB:

Qy	753	SerAspIlePheArgValAlaIleAlaGlyAlaProValThrLeuThrIlePheTyrAsp	772	Qy	115	SerTrpLysProLeuLeuAspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArg	134
Db	2596	CCCAGGTGTTCAAGGTGCCATCGCGGGTCCCGGTCACCGTCTGGATCGCCTACGAC	2655	Db	676	TCTTGGAGAGAGATGCTGGATCATTTTCCAGGCCACGCCACCATGGGTCTACTCTCGG	735
Qy	773	ThrGlyTyrThrGluArgTyrMetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeu	792	Qy	135	GluGluGluLeuLeuArgGluArgLysArgIleGlyThrValGlyLleAlaSerTyrAsp	154
Db	2656	ACAGGGTACACTGAGCGCTACATGGACGTCCTCCCTGGAGAACACACGACGCGCTATGAGCGG	2715	Db	736	GAGGAGGAGCTGCTGAGGAGAGCGAAACGCTGGGGGTCTTGGCATCACCTCTACGAC	795
Qy	793	GlySerValAlaMetGlnAlaGluLysPheProSerGluProAsnArgLeuLeuLeu	812	Qy	155	TyrHisGlnGlySerGlyThrPheLeuPheGlnAlaGlySerGlyLleTyrHisValLys	174
Db	2716	GTTTCCGTGGCCCTGCACGTCGGAGAAAGCTGCCCAATGAGCCCAACCGCTTGTCTATCCTC	2775	Db	796	TTCCACAGCAGAGTGGCTCTTCTCTTCCAGGCCACGACACAGCCTCTTCCACTCCGCG	855
Qy	813	HisGlyPheLeuAspGluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeu	832	Qy	175	AspGlyGlyProGlnGlyPheThrGlnGlnProLeuArgProAsnLeuValGluThrSer	194
Db	2776	CACGGCTTCTGGACGAAACGTCGCACTTTTTCACACAAACTTCTCTGCTCTCCCACTG	2835	Db	856	GACGGCGCCCGATGGACCCCAAAATCTGCCCTGACGACCTCGTAATCAAGACCCAG	915
Qy	833	ValArgAlaGlyLysProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleArg	852	Qy	195	CysProAsnIleArgMetAspProLysLeuCysProAlaAspProAspTrpIleAlaPhe	214
Db	2836	ATCCGAGCAGGGAAACCTTACCAGCTCCAGATCTACCCCAACGAGACACAGATATTGCG	2895	Db	916	TGCTCAGGGCCCGGATGGACCCCAAAATCTGCCCTGACGACCTCTTCTCTCTCTC	975
Qy	853	ValProGluSerGlyGluHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeu	872	Qy	215	IleHisSerAsnAspIleTrpIleSerAsnIleValThrArgGluGluArgLeuThr	234
Db	2896	TGCCCCGAGTCGGGAGCACTATGAAGTCACGTTGCTGCACTTTCTACAGGAATACCTC	2955	Db	976	ATCAATAACAGCAGCCTGTGGTGGCCAAACATCGACACAGGCGAGGCGGCGGTGACC	1035
RESULT 11							
US-09-976-674-24							
; Sequence 24, Application US/09976674							
; Patent No. 6844180							
; GENERAL INFORMATION:							
; APPLICANT: Qi, Steve							
; APPLICANT: Akinsanya, Karen							
; APPLICANT: Riviere, Pierre							
; APPLICANT: Junien, Jean-Louis							
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV							
; FILE REFERENCE: 70669							
; CURRENT APPLICATION NUMBER: US/09/976,674							
; PRIOR FILING DATE: 2001-10-12							
; PRIOR APPLICATION NUMBER: US 60/240,117							
; PRIOR FILING DATE: 2000-10-12							
; NUMBER OF SEQ ID NOS: 61							
; SOFTWARE: PatentIn version 3.1							
; SEQ ID NO 24							
; LENGTH: 4302							
; TYPE: DNA							
; ORGANISM: Homo sapiens							
US-09-976-674-24							
Alignment Scores:							
Pred. No.: 4,48e-315 Length: 4302							
Score: 2670.00 Matches: 517							
Percent Similarity: 77.5% Conservative: 134							
Best Local Similarity: 61.5% Mismatches: 187							
Query Match: 61.1% Indels: 2							
DB: 3 Gaps: 2							
US-10-825-632-1 (1-882) x US-09-976-674-24 (1-4302)							
Qy	35	PheTyrValGluArgTyrSerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLys	54	Qy	394	GluLeuPheIleProValGluAspValMetGluArgGlnArgLeuIleGluSerVal	413
Db	436	TTCAGGTGCAAGACACTGCTGGGACCGGCTCCGGAGCATCATCCAGCGCACCCGAC	495	Db	1516	GCCCTGTTTCATCCGAGCACAGAGATGAGGAGCAGCGGCTAGCCTCTGCCAGAGCTGC	1575
Qy	55	TyrHisGlyTyrMetMetAlaLysAlaProHisAspPheMetPheValLysArgAsnAsp	74	Qy	414	ProAspSerValThrProLeuIleIleTyrGluGluThrThrAspIleTrpIleAsnIle	433
Db	496	TACTCGGGCTCATTTGTCAACAGCGGCCCAACACTTCCAGTTTGTGCAAGACCGAT	555	Db	1576	CCAGGAATGTCAGCGGTATGTGGTGACGAGGAGGTCAACACGCTGGATCAATGTT	1635
Qy	75	ProAspGlyProHisSerAspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGlu	94	Qy	434	HisAspIlePheHisValPheProGlnSerHis---GluGluGluIleGluPheIlePhe	452
Db	556	GAGTCTGGGCGCCCACTCCACCGCTCTACTACTTGGGAATGCCATATGCCAGCCGAG	615	Db	1636	CATGACATCTTCTCCCTTCCCAATCAGAGGAGGAGGACGAGCTCTGCTTCTCCGC	1695
Qy	95	AsnThrLeuPheTyrSerGluIleProLysThrIleAsnArgAlaAlaValLeuMetLeu	114	Qy	453	AlaSerGluCysLysThrGlyPheArgHisLeuTyrLysValIleThrSerIleLeuLysGlu	472
Db	616	AACCTCCCTCTCTACTCTGAGATTCCCAAGAGGTCCGGAAGAGGCTCTGCTGCTCCTG	675	Db	1696	GCCATGAAATGCAAGACCGGCTTCTGCCATTGTACAAGTCAACCGCGCTTTAAATCC	1755
Qy				Qy	473	SerLysTyrIleArgSerSerGlyGlyLeuProAlaProSerAspPheLysCysProIle	492

Db 1756 CAGGGCTACGATTGAGTGGCCCTTCAGCCCGGGGAGAGATGAATTAAGTGCCCATTT 1815
Qy LysGluGluLeuAlaIleThrSerGlyGluTrpGluValLeuGlyArgHisGlySerAsn 512
Db 1816 AAGGAAGAGATTGCTCTGACACACGGTGAATGGAGAGTTTGGCGAGGACGGCTCCAG 1875
Qy 513 IleGlnValAspGluValArgArgLeuValTyrPheGluGlyThrLysAspSerProLeu 532
Db 1876 ATCTGGGTCAATGAGGAGACCAAGCTGGTGTACTTCCAGGGCCACCAAGGACACGCGCTG 1935
Qy 533 GluHisThrLeuValValSerTyrValAsnProGlyGluValThrArgLeuThrAsp 552
Db 1936 GAGCACCACCTCTACGTGGTGTGAGCTATGAGCGCGCGCGGAGATCGTAGCGCTCACCA 1995
Qy 553 ArgGlyTyrSerHisSerCysValIleSerGlnHisCysAspPhePheIleSerLysTyr 572
Db 1996 CCGCGCTTCTCCCACTAGCTGCTCATGAGCCAGCAACTTCGACATGTTGTGACGCACTAC 2055
Qy 573 SerAsnGlnLysAsnProHisCysValSerLeuTyrLysLeuSerSerProGluAsp 592
Db 2056 AGCAGCGTGAGCAGCGCGCTGCTGCACTCTCAAGCTGAGCGCGCGCGGAGCAGCAG 2115
Qy 593 ProThrCysLysThrLysGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuPro 612
Db 2116 CCGCTGCACAGCAGCCCGCTTCTGGCTGTAGCATGATGAGCGCAGCCAGCTGCCCGCG 2175
Qy 613 AspTyrThrProProGluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGly 632
Db 2176 GATTATGTTCCTCAGAGATCTTCATTTTCCACACCGCTCGGATGTGCGGCTCTACGGC 2235
Qy 633 MetLeuTyrLysProHisAspLeuGlnProGlyLysLysTyrProThrValLeuPheIle 652
Db 2236 ATGATCTACAGCCCGCGCTGCGCGCAGGAGAGCAGCAGCCACCGCTCTCTTGTGA 2295
Qy 653 TyrGlyGlyProGlnValGlnLeuValAsnAsnArgPheLysGlyValLysTyrPheArg 672
Db 2296 TATGGAGGCGCCCGAGGTGCAGTGGTGAATACTCTTCAAGGCGATCAAGTACTTGGCG 2355
Qy 673 LeuAsnThrLeuAlaSerLeuGlyTyrValValValValIleAspAsnArgGlySerCys 692
Db 2356 CTCACACACTGGCTCCCTCGGCTAGCGCTGGTGTGATGACGCGGCGGCTCTGTGT 2415
Qy 693 HisArgGlyLeuLysPheGluGlyAlaPheLysTyrLysMetGlyGlnIleGluLeuAsp 712
Db 2416 CAGCAGGCGCTTCGGTTCGAAAGGCGCTCGAAAAACCAATGGCCAGGTGGAGATCGAG 2475
Qy 713 AspGlnValGluGlyLeuGlnTyrLeuAlaSerArgTyrAspPheIleAspLeuAspArg 732
Db 2476 GACCAGGTGGAGGCGCTGCAGTTGCTGGCGGAGAGATGATGGCTTCATCGACCTGAGCGGA 2535
Qy 733 ValGlyIleHisGlyTyrSerTyrGlyTyrLysSerLeuMetAlaLeuMetGlnArg 752
Db 2536 GTTGCATTCATGCGTGGCTCTAGCGGGGCTCTCTCGCTCATGCGGGGTAAATCCACAG 2595
Qy 753 SerAspIlePheArgValAlaIleAlaGlyAlaProValThrLeuTrpIlePheTyrAsp 772
Db 2596 CCCCAGGTGTTCAAGGTGGCCATCGCGGGTGGCCCGGTACCGCTCTGGATGGCTTACGAC 2655
Qy 773 ThrGlyTyrThrGluAlaTyrMetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeu 792
Db 2656 ACAGGGTACATGAGCGCTACATGAGCGTCCCTGAGNACACACCGAGCAGCGCTATGAGCG 2715
Qy 793 GlySerValAlaMetGlnAlaGluLysPheProSerGluProAsnArgLeuLeuLeu 812
Db 2716 GGTTCGGTGGCGCTGCAGTGGAGAGCTGCCCAATGAGCCCAACCGCTTGTCTTATCCTC 2775
Qy 813 HisGlyPheLeuAspGluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeu 832
Db 2776 CACGGCTTCTTGAGAGAAACGTCACCTTTTTCACACAAACTTCTCGTCTCCCACTG 2835
Qy 833 ValArgAlaGlyLysProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleArg 852

Db 2836 ATCCGAGCAGGGAACCTTACCAGCTCCAGATCTACCCCAACGAGACACAGTATTCCG 2895
Qy 853 ValProGluSerGlyGluHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeu 872
Db 2896 TGCCCGGAGTGGCGGAGCACTATGAAGTCAAGTGTGTCGACTTTCTACAGGAATACCTC 2955
RESULT 12
US-09-976-674-36
; Sequence 36, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 36
; LENGTH: 4180
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-976-674-36
Alignment Scores:
Pred. No.: 1,82e-309 Length: 4180
Score: 2820.50 Matches: 510
Percent Similarity: 76.4% Conservative: 132
Best Local Similarity: 60.7% Mismatches: 183
Query Match: 60.0% Indels: 15
Gaps: 3
US-10-825-632-1 (1-882) x US-09-976-674-36 (1-4180)
Qy 35 PheTyrValGluArgTyrSerTyrSerGlnLeuLysLysLeuLeuAlaAspThrArgLys 54
Db 436 TTCAGGTGTCAGAGACATCTGTGGACGGGTCCGGAGCATCATCCACGCGCAGCGCAG 495
Qy 55 TyrHisGlyTyrMetMetAlaLysAlaProHisAspPheMetPheValLysArgAsnAsp 74
Db 496 TACTCGGCGCTCATTTGTCACAAAGCGCGCCCGCCAGCTTCCAGTTTGTGCAGAAAGCGAT 555
Qy 75 ProAspGlyProHisSerAspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGlu 94
Db 556 GAGTCTGGGCGCCACTCCCGCTCTACTACCTGGGAAATGCCATATGCGAGCGCGAGAG 615
Qy 95 AsnThrLeuPheTyrSerGluIleProLysThrIleAsnArgAlaAlaValLeuMetLeu 114
Db 616 AACTCCCTCTCTACTCTGAGATTCGAGAGGTCCGGAAGAGGCTCTGCTGCTCTG 675
Qy 115 SerTrpLysProLeuAspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArg 134
Db 676 TCTTGAAGCAGATGCTGGATCATTTCCAGCCAGCGCCACCATCTCTCTCTCTCGG 735
Qy 135 GluGluGluLeuLeuArgGluArgLysArgIleGlyThrValGlyIleAlaSerTyrAsp 154
Db 736 GAGGAGGAGCTCTGAGGAGCGGAAACGCTGGGGGTCTTTCGGCATCATCTCTCTACGAC 795
Qy 155 TyrHisGlnGlySerGlyThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLys 174
Db 796 TTCACAGGAGAGTGGCGCTCTTCTCTTCCAGGCGGAGCAACACCTCTTCCACTGCGCG 855
Qy 175 AspGlyGlyProGlnGlyPheThrGlnGlnProLeuArgProAsnLeuValGluThrSer 194
Db 856 GACGGGCGGAGAAACGCGCTTCTATGCTGCTCCCTATGAAACCGCTGGAAATCAAGACCCAG 915
Qy 195 CysProAsnIleArgMetAspProLysLeuCysProAlaAspProAspTrpIleAlaPhe 214

Db 916 TGCTAGGCGCCGATGACCCCAAAATCTGCCCTCGCGACCCCTGCTTCTTCTCTC 975
Qy 215 ILeHisSerAenAspIleThrPheSerAenIleValThrArgGluGluArgLeuThr 234
Db 976 ATCAATAACAGCGACCTGTGGTGGTCAACATCTGACAGACGAGCGGCGCTGACC 1035
Qy 235 TyrValHisAenGluLeuAlaAenMetGluGluAaspAlaArgSerAlaGluValAlaThr 254
Db 1036 TTCTGCCACCAAGGTTATCAATGTCCTGGATGACCCCAAGTCTGGGGTGTGGCCACC 1095
Qy 255 PheValLeuGluGluPheAepArgTyrSerGlyTyrTTPTrpCysPheProLeuAlaGlu 274
Db 1096 TTCTGTATACAGAAAGTTCCAGCCGCTTCACTGGGTACTGGTGGTCCCCACAGCCTCC 1155
Qy 275 ThrThrProSerGlyGly---LysIleLeuArgIleLeuTyrGluGluAenAspGluSer 293
Db 1156 TGGAAAGGTTACAGAGGCTCAAGACGCTGCGAATCTGTATGAGGAAGTCCATGAGTCC 1215
Qy 294 GluValGluIleIleHisValThrSerProMetLeuGluThrArgArgAlaAaspSerPhe 313
Db 1216 GAGGTGAGGTCAATCACGTCCTCTCTGCGCTAGAAAGAAAGACCGGACTCGTAT 1275
Qy 314 ArgTyrProLysThrGlyThrAlaAenProLysValThrPheLysMetSerGluIleMet 333
Db 1276 CGGTACCCAGGACAGCAGCAAGAAATCCCAAGATTGCCCTGAAACTGGCTGAGTCCAG 1335
Qy 334 IleAaspAlaGluGlyArgIleLeuAaspValIleAaspLysGluLeuIleGlnProPheGlu 353
Db 1336 ACTGACAGCCAGGCGAAGATCGTCTCGACCCAGGAGAGGAGCTGGTCAGCCCTTCAGC 1395
Qy 354 IleLeuPheGluGlyValGluTyrIleAlaArgAlaGlyTyrThrProGluGlyLysTyr 373
Db 1396 TCGCTGTTCCGAGGTGGAGTACATCGCCAGGCGCGGTGGACCCGGGTGCAAAATAC 1455
Qy 374 AlaTrpSerIleLeuAenAaspArgSerGlnThrArgLeuGlnIleValLeuIleSerPro 393
Db 1456 GCTTGGGCGCATGTCCTGGACCGCCGCCAGCGTCCAGCTCGTCTCTCTCCCGCG 1515
Qy 394 GluLeuPheIleProValGluAaspAvalMetGluArgGlnArgLeuIleGluSerVal 413
Db 1516 GCCCTGTTCATCCCGACACAGAGATGAGGACAGCGCTAGCCTTGCACAGAGTCTC 1575
Qy 414 ProAaspSerValThrProLeuIleIleTyrGluGluThrThrAaspIleThrIleAenIle 433
Db 1576 CCCAGGAATGTCACGCGGTATGTGGTGTACGAGGAGGTCAACAGCTCTGGATCAATGT 1635
Qy 434 HisAepIlePheHisValPheProGlnSerHis---GluGluGluIleGluPheIlePhe 452
Db 1636 CATGACATCTTCTATCCCTTCCCCCAATCAGAGGAGAGGAGAGCTCTGCTTCTCCGC 1695
Qy 453 AlaserGluCysLysThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLysGlu 472
Db 1696 GCCAATGAATCAAGACCGGCTTCTGCCATTTGTCAAAAGTCAACCGCCGCTTTTAAATCC 1755
Qy 473 SerLysTyrLysArgSerSerGlyLeuProAlaProSerAaspPheLysCysProIle 492
Db 1756 CAGGCTACGATTGAGTGGCTTCCAGCCCGCGGGAAGATGAATTTAAGTGCCTCAT 1815
Qy 493 LysGluGluIleAlaIleThrSerGlyGluTyrGluValLeuGluArgHisGlySerAen 512
Db 1816 AAGGAAGATGCTCTGACACCGGCTGAATGGAGGTTTGGCAGGACCGGCTCC--- 1872
Qy 513 IleGlnValAaspGluValArgArgLeuValTyrPheGluGlyThrLysAaspSerProLeu 532
Db 1873 -----AAGGGCACCAAGACACGCGCTG 1896
Qy 533 GluHisLeuLeuTyrValValSerTyrValAenProGlyGluValThrArgLeuThrAasp 552
Db 1897 GAGCAACCACTCTACGTGGTCACTATGAGCGCGCGCGAGATCGTACGCTTCACACG 1956
Qy 553 ArgGlyTyrSerHisSerCysCysIleSerGlnHisCysAaspPhePheLeuSerLysTyr 572
Db 1957 CCCGGCTTCTCCCATAGCTGCTCCATGAGCCAGAACTTCGACATGTTCGTGACCCACTAC 2016

Qy 573 SerAenGlnLysAenProHisCysValSerLeuTyrLysLeuSerSerProGluAasp 592
Db 2017 AGCAGCTGACGACCGCCCTCGCTGCGAGCTTACAGCTAGCGCGCCCGACGACGAC 2076
Qy 593 ProThrCysLysThrLysGluPheTrpAlaThrIleLeuAaspSerAlaGlyProLeuPro 612
Db 2077 CCCCTGCACAAAGCAGCCCGCTTCTGGGTAGCATGTGAGGACGACGAGTCCGCCCG 2136
Qy 613 AspTyrThrProProGluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGly 632
Db 2137 GATTATGTTCTCCAGAGATCTTCCATTTCCACACCGCTCGGATGTGCGGCTCTACGGC 2196
Qy 633 MetLeuTyrLysProHisAaspLeuGlnProGlyLysLysTyrProThrValLeuPheIle 652
Db 2197 ATGATCTACAAAGCCCGCTTGCAGCCAGGAGGAGAACGCCCCACCGCTCTCTTTGTA 2256
Qy 653 TyrGlyGlyProGlnValGlnLeuValAenAenArgPheLysGlyValLysTyrPheArg 672
Db 2257 TATGGAGCCCGCAGGTGCAGTGGTGAATACTCTTCAAGGCATCAAGTACTTGGCG 2316
Qy 673 LeuAenThrLeuAlaSerLeuGlyTyrValValValIleAaspAenArgGlySerCys 692
Db 2317 CTCACACACTGGCTCCCTGGCTAGCCGCTGTGTGATTGACGCGAGGGCTCTCTGT 2376
Qy 693 HisArgGlyLeuLysPheGluGlyAlaPheLysTyrLysMetGlyGlnIleGluLeuAap 712
Db 2377 CAGCGAGGCTTCGGTTCCGAAGGGCCCTCGAAACCAATGGGCCAGGTGGAGATCGAG 2436
Qy 713 AepGlnValGluGlyLeuGlnTyrLeuAaserArgTyrAaspPheIleAaspLeuAepArg 732
Db 2437 GACCAAGTGGAGGCGCTGCAGTTCGTGGCGGAGAAAGTATGGCTTCATCGACCTGACCGA 2496
Qy 733 ValGlyIleHisGlyTyrSerTyrGlyTyrLeuSerLeuMetAlaLeuMetGlnArg 752
Db 2497 GTTGCATCTCAGCTGGTCTCTACGGGGGCTTCTCTCGCTCATGGGGCTTAATCCACAAG 2556
Qy 753 SerAaspIlePheArgValAlaIleAlaGlyAalaproValThrLeuTrpIlePheTyrAap 772
Db 2557 CCCAGGTTTCAAGGTGGCCATCGCGGGTCCCGGCTCACCGCTCTGGATGGGCTACGAC 2616
Qy 773 ThrGlyTyrThrGluArgTyrMetGlyHisProAaspGlnAasnGluGlnGlyTyrLeu 792
Db 2617 ACAGGGTACACTGACGCTACATGAGCGTCCCTGAGAAACAACGACGCGCTATGAGGCG 2676
Qy 793 GlySerValAlaMetGlnAlaGluLysPheProSerGluProAenArgLeuLeuLeuLeu 812
Db 2677 GGTTCGTGGCCCTGCAGCTGAGAGAGTCCCAATGAGCCCAACCGCTTGTCTTATCCTC 2736
Qy 813 HisGlyPheLeuAaspGluAenValHisPheAlaHisThrSerIleLeuLeuSerPheLeu 832
Db 2737 CACGGCTTCTGGAGGAAACGTGCATTTTCCACCAAACTTCTCTGCTCTCCCACTG 2796
Qy 833 ValArgAlaGlyLysProTyrAaspLeuGlnIleTyrProGlnGluArgHisSerIleArg 852
Db 2797 ATCCGAGCAGGGAACCTTACAGCTCCAGATCTACCCCAACGAGACACAGTATTCCG 2856
Qy 853 ValProGluSerGlyGluHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGluAenLeu 872
Db 2857 TGGCCCGAGTGGGCGGAGCACTATGAAGTCACTGTGTGCACTTTTACAGGAATACCTC 2916

RESULT 13
US-09-976-674-34
; Sequence 34, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976, 674

Tue Apr 18 08:18:33 2006

```

; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 34
; LENGTH: 4263
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-976-674-34

```

```

Alignment Scores:      1.88e-309      4263
Pred. No.:            2820.50      510
Score:                 76.4%      132
Percent Similarity:    60.7%      183
Best Local Similarity: 60.0%      15
Query Match:           3          3
DB:                    3          3

```

US-10-825-632-1 (1-882) x US-09-976-674-34 (1-4263)

QY	35	PheTyrValGluArgTyrSerTrpSerGlnLeuLysLeuLeuAlaAspThrArgLys	54
DB	436	TTCCAGGTGCAGAGCACTCGTGGACGGGCTCCGAGCATCATCCACGGCAGCGCAAG	495
QY	55	TyrHisGlyTyrMetMetAlaLysAlaProHisAspPheMetPheValLysArgAsnAsp	74
DB	496	TACTCGGGCCTCATTTGTCACAGGCGCCCGACACTTCCAGTTTGTGCAAGACCGAT	555
QY	75	ProAspGlyProHisSerAspArgIleTyrTyrLeuAlaMetSerGlyLysAsnArgGlu	94
DB	556	GAGTCTGGGCCCCACTCCACGGCTCTACTCTGGGAATGCCATATGGCAGCGGAGAG	615
QY	95	AsnThrLeuPheTyrSerGluIleProLysThrIleAsnArgAlaAlaValLeuMetLeu	114
DB	616	AACTCCCTCTCTACTCTGAGATTCCCAAGAGTCCGGAAGAGGCTCTGCTGCTCTG	675
QY	115	SerTrpLysProLeuLeuAspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArg	134
DB	676	TCCTGGAAGACGATGTGGATCAVTTCCAGGCCACGCCCCACCATGGGGGTCTACTCTCGG	735
QY	135	GluGluGluLeuLeuArgGluArgLysArgIleGlyThrValGlyIleAlaSerTyrAsp	154
DB	736	GAGGAGGAGCTGCTGCGGGAGCGGAACGGCTGGGGGTCTTGGCATCACCTCTCTACGAC	795
QY	155	TyrHisGlnGlySerGlyThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLys	174
DB	796	TTCCACAGCGAGAGTGGGCTCTTCTCTTCCAGGCCACCAAGAGCTCTTCCACTTGGCGC	855
QY	175	AspGlyGlyProGlnGlyPheThrGlnGlnProLeuArgProAsnLeuValGluThrSer	194
DB	856	GACGGCGGCAGAACCGGCTTCATGGTGTCCCTTATGAACCGCTGGAATCAAGACCCAG	915
QY	195	CysProAsnIleArgMetAspProLysLeuCysProAlaAspProAspTrpIleAlaPhe	214
DB	916	TGCTCAGGGCCCCGATGACCCCAAAATCTGCGCTCGCGACCCCTCTCTCTCTCTC	975
QY	215	IleHisSerAsnAspIleTrpIleSerAsnIleValThrArgGluGluArgGluThr	234
DB	976	ATCAATAACAGCGACCTGTGGGTGGCCCAACATCGACAGCGGAGGAGCGGGGTGACC	1035
QY	235	TyrValHisAsnGluLeuAlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThr	254
DB	1036	TTCTGCGCCACCAAGGTTTATCCAAATGTCTCGATGACCCCAAGTCTGGCGGGTGTGCCACC	1095
QY	255	PheValLeuGlnGluPheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGlu	274
DB	1096	TTGCTCATACAGAAAGATTGACCGCTTCACTGGGTACTTGTGTGTGGTGGCCACGCTCC	1155
QY	275	ThrThrProSerGlyGly---LysIleLeuArgIleLeuTyrGluGluAsnAspGluSer	293
DB	1156	TGGGAGGTTTCAGAGGGCTTCAGACGCTCGAATCTGTATGAGGAAGTCAATGAGTCC	1215

QY	95	AsnThrLeuPheTyrSerGluIleProLysThrIleAsnArgAlaValLeuMetLeu	114
Db	616	AACTCCCTCTCTACTCTGAGATTCCCAAGAGGTCCGGAAAGAGGCTCTGCTCTCTG	675
QY	115	SerTrpLysProLeuLeuAspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArg	134
Db	676	TCCTGGAGACAGATGCTGGATCATTCCTCCAGGCCACGCCCCACCACTGGGGTCTACTCTCGG	735
QY	135	GluGluGluLeuLeuArgGluArgLysArgIleGlyThrValGlyIleAlaSerTyrAsp	154
Db	736	GAGGAGAGCTGCTGAGGGACGGGAACGCTCGGGCTCTCGGCATCACCTCTCAGCAC	795
QY	155	TyrHisGlnGlySerGlyThrPheLeupheGlnAlaGlySerGlyIleTyrHisValLys	174
Db	796	TTCCACACGAGAGTGGCTCTTCCTCTCCAGGCCAGCAACAGCCTCTCCACTCGCGC	855
QY	175	AspGlyGlyProGlnGlyPheThrGlnGlnProLeuArgProAsnLeuValGluThrSer	194
Db	856	GACGGCGCAAGACGGCTTCATGGTGTCCCTATGAACCTCGAATCAAGACCCAG	915
QY	195	CysProAsnIleArgMetAspProLysLeuCysProAlaAspProAspTrpIleAlaPhe	214
Db	916	TGCTCAGGCGCCCGGATGGACCCCAAAATCTGCCCTGCGACCTCTCTTCCTTC	975
QY	215	IleHisSerAsnAspIleTrpIleSerAsnIleValThrArgGluGluArgLeuThr	234
Db	976	ATCAATAACAGCGACTGTGGGTGGCCAAACATCGACACAGCGGAGAGCGCGCTGACC	1035
QY	235	TyrValHisAsnGluLeuAlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThr	254
Db	1036	TTCTGCCACCAAGGTTTATCCAATGTCTCGATGACCCCAAGTCTCGGGGTGGCCACC	1095
QY	255	PheValLeuGlnGluGluPheAspArgTyrSerGlyTyrTrpCysProLysAlaGlu	274
Db	1096	TTCTGTCATACAGAAAGATTTCGACGGTTCATCGGTACTGGTGTGGTGCCTCCACAGCCTCC	1155
QY	275	ThrThrProSerGlyGlyLysIleLeuArgIleLeuTyrGluLeuAsnAspGluSer	293
Db	1156	TGGGAAGTTTCAGGGGCTCAAGCGCTCGAATCTCTGTATGAGGAAGTCGATGATGCC	1215
QY	294	GluValGluIleIleHisValThrSerProMetLeuGluThrArgAlaAspPhe	313
Db	1216	GAGTGGAGGTCAATTCAGTCCCTCTCTCCGCTAGAGAAAGGAGCGGACTCGTAT	1275
QY	314	ArgTyrProLysThrGlyThrAlaAsnProLysValThrPheLysMetSerGluIleMet	333
Db	1276	CGGTACCCAGGACAGGACGAGCAAGAAATCCCAAGATGCTTGAACCTGCTGATGCCAG	1335
QY	334	IleAspAlaGluGlyArgIleLeuAspValIleAspLysGluIleLeuGlnProPheGlu	353
Db	1336	ACTGACAGCCAGGCGAAGATCGTCTCGACCCAGGAGAGGAGCTGGTGGCCCTTCAGC	1395
QY	354	IleLeuPheGluGlyValGluTyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyr	373
Db	1396	TCGCTGTTCGGAAGTGGAGTACATCGCCAGGCGCGGGTGGACCGGATGGCAATATC	1455
QY	374	AlaTrpSerIleLeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSerPro	393
Db	1456	GCCTGGGCCATGTCTCTGACCGGCCCCAGCAGTGGCTCCAGCTCGCTCTCCCTCCCGCG	1515
QY	394	GluLeuPheIleProValGluAspValMetGluArgGlnArgLeuIleGluSerVal	413
Db	1516	GCCTGTTCATCCCGAGCACAGAGAATGAGGACGCGCTAGCCTCTGCGCAGAGTGTCT	1575
QY	414	ProAspSerValThrProLeuIleIleTyrGluGluThrThrAspIleTrpIleAsnIle	433
Db	1576	CCNAGGAATCTCAGCGGTATGTGGTGTACGAGGAGGTCCACCAAGCTCTGGATCAATGT	1635
QY	434	HisAspIlePheHisValPheProGlnSerHisGluGluGluIleGluPheIlePhe	452
Db	1636	CATGACATCTCTATCCCTTCCCCCAATCAGAGGAGGAGGACGAGCTCTGCTTCTCCCG	1695
QY	453	AlaSerGluCysLysThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLysGlu	472
Db	1696	GCCATGATGCAAGACCGGCTTCGCCATTTGTACAAAGTCAACGCGCTTTAAATCC	1755
QY	473	SerLysTyrLysArgSerSerGlyGlyLeuProAlaProSerAspPheLysCysProIle	492
Db	1756	CAGGGCTACGATGGAGTGGAGCCCTTCAGCCCGGGGAAGATGAATTAAGTCCCAT	1815
QY	493	LysGluGluIleAlaIleThrSerGlyGluTrpGluValLeuGlyArgHisGlySerAsn	512
Db	1816	AAGAAAGAGATGTGCTCTGACCGGTGAATGGAGGTTTTTGGCAGCACCGCTCCAAG	1875
QY	513	IleGlnValAspGluValArgArgLeuValTyrPheGluGlyThrLysAspSerProLeu	532
Db	1876	ATCTGGGTCAATGAGGAGACCAACAGCTGGTGATCTTCAGGAGCACCAAGACACGCGCTG	1935
QY	533	GluHisHisLeuTyrValValSerTyrValAsnProGlyGluValThrArgLeuThrAsp	552
Db	1936	GAGCACCACTCTACGTGGTCACTATGAGCGCGCGCGGAGATCGTAGCGCTCACACG	1995
QY	553	ArgGlyTyrSerHisSerCysCysIleSerGlnHisCysAspPhePheIleSerLysTyr	572
Db	1996	CCCGCTTCTCCCATAGCTCTCCATGAGCCAGAACTTCGACATGTTCGTCAGCCACTAC	2055
QY	573	SerAsnGlnLysAsnProHisCysValSerLeuTyrLysLeuSerSerProGluAspAsp	592
Db	2056	AGCAGCGTGACACGCGCGCTGCTGCTCAAGCTGAGCGCGCCCGACGACGAC	2115
QY	593	ProThrCysLysThrLysGluPheTyrAlaThrIleLeuAspSerAlaGlyProLeuPro	612
Db	2116	CCCTGTCACAGACCGCCGCTTCGGCTAGCATGTGAGGCGCAGCCAGCTGCCCGCG	2175
QY	613	AspTyrThrProProGluIlePheSerPheGluSerThrGlyPheThrLeuTyrGly	632
Db	2176	GATTATGTCTCCAGAGATCTTCATTTCCACGCGCTCGATGCGGCTCTACGCG	2235
QY	633	MetLeuTyrLysProHisAspLeuGlnProGlyLysLysTyrProThrValLeuPheIle	652
Db	2236	ATGATCTCAAGCCCGCCGCTTCGACCGAGGAGAGCCGCCACCTCTCTCTTGTGA	2295
QY	653	TyrGlyGlyProGlnValGlnLeuValAsnAsnArgPheLysGlyLysLysTyrPheArg	672
Db	2296	TATGGAGCGCCCGAGTGCAGTGGTGAATAACTCTTCAAGGCATCAAGTACTTGC	2355
QY	673	LeuAsnThrLeuAlaSerLeuGlyTyrValValValIleAspAsnArgGlySerCys	692
Db	2356	CTCAACACACTGGCTCTCGGCTACGCGTGTGTGTGATTGACGCGCGGCTCTCTGT	2415
QY	693	HisArgGlyLeuLysPheGluGlyAlaPheLysTyrLysMetGlyGlnIleGluIleAsp	712
Db	2416	CAGCGAGGCTTCGGTTCGAAGGGGCTCGAAAAACCAATGGGCCAGGTGGAGATCGAG	2475
QY	713	AspGlnValGluGlyLeuGlnTyrLeuAlaSerArgTyrAspPheIleAspLeuAspArg	732
Db	2476	GACCACTGGAGGCGCTTCAGTTCGTGGCCGAGAAGTATGCTTCATCGACCTGAGCCGA	2535
QY	733	ValGlyIleHisGlyTrpSerTyrGlyTyrLeuSerLeuMetAlaLeuMetGlnArg	752
Db	2536	GTTGCCATCTCATGGCTGGTCTCTACGGGGCTCTCTCGCTCATGGGGCTTAATCCACAAG	2595
QY	753	SerAspIlePheArgValAlaIleAlaGlyAlaProValThrLeuTrpIlePheTyrAs	772
Db	2596	CCCAAGTGTTCAGG-----	2611
QY	772	pThrGlyTyrThrGluArgTyrMetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLe	792
Db	2611	-----	2611
QY	792	uGlySerValAlaMetGlnAlaGluLysPheProSerGluProAsnArgLeuLeuLe	812
Db	2612	-----CCCAACCGCTGCTTATCCT	2631
QY	812	uHisGlyPheLeuAspGluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLe	832

Db	2632	CCACGGCTTCCTGGACGAAACGTGCACCTTTTCCACACAACTTCTCGTCTCCCACT	2691
QY	832	uValArgAlaGlyLysProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleAr	852
Db	2692	GATCCGACGAGGGAACCTTACCCAGCTCCAGATCTACCCCAACGAGACACAGTATTCG	2751
QY	852	gValProGluSerGlyGluHisTyrGluHisLeuLeuHisTyrLeuGlnGluAsnLe	872
Db	2752	CTGCCCCGAGTCGGGCGAGCAGCTATGAAGTCAGTTGCTGCACCTTTCTACAGGAATACCT	2811
QY	872	u 872	
Db	2812	C 2812	

Search completed: April 15, 2006, 01:02:11
Job time : 495 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model
Run on: April 15, 2006, 01:02:29 ; Search time 1458 Seconds
(without alignments)
5002.463 Million cell updates/sec

Title: US-10-825-632-1
Perfect score: 4700
Sequence: 1 MAAMETEQLGVETETADC.....HLLHYLQENLGSRIALAKVI 882

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 9793542 seqs, 4134689005 residues
Total number of hits satisfying chosen parameters: 19587084
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame_plus_p2n.model -DEV=xlp
-Q-/abses/ABSSWEB spool/US10825632/runat 14042006 092611 4971/app query.fasta_1
-DB=Published Applications NA_Main -QFMT=fastap -SUFFIX=p2n.rnpbm
-MINMATCH=0.1 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct
-THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext
-HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abes05p
-USER=US10825632 @CGN 1 1 1549 @runat 14042006 092611 4971 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA_Main:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4700	100.0	2649	US-10-054-776-1	Sequence 1, Appl
2	4700	100.0	2649	US-10-170-789-39	Sequence 39, Appl
3	4700	100.0	2671	US-09-976-674-2	Sequence 2, Appl
4	4700	100.0	2671	US-10-982-512-2	Sequence 2, Appl
5	4700	100.0	3106	US-10-311-035-30	Sequence 30, Appl
6	4700	100.0	3120	US-10-415-122-5	Sequence 5, Appl
7	4700	100.0	3120	US-10-825-632-2	Sequence 2, Appl

8	4700	100.0	3143	6	US-10-170-789-37	Sequence 37, Appl
9	4680	99.6	4829	3	US-09-976-674-12	Sequence 12, Appl
10	4680	99.6	4829	9	US-10-982-512-12	Sequence 12, Appl
11	4385.5	93.3	4685	3	US-09-976-674-22	Sequence 22, Appl
12	4385.5	93.3	4685	9	US-10-982-512-22	Sequence 22, Appl
13	4385	93.3	4676	3	US-09-976-674-20	Sequence 20, Appl
14	4385	93.3	4676	9	US-10-982-512-20	Sequence 20, Appl
15	4118	87.6	2830	3	US-10-956-157-2177	Sequence 2177, Ap
16	4092.5	87.1	4523	3	US-09-976-674-8	Sequence 8, Appl
17	4092.5	87.1	4523	9	US-10-982-512-8	Sequence 8, Appl
18	3970.5	84.5	2510	7	US-10-275-505-16	Sequence 16, Appl
19	3970.5	84.5	2510	10	US-11-140-224-16	Sequence 16, Appl
20	3661.5	77.9	4309	3	US-09-976-674-14	Sequence 14, Appl
21	3661.5	77.9	4309	9	US-10-982-512-14	Sequence 14, Appl
22	2870	61.1	2617	3	US-09-976-674-4	Sequence 4, Appl
23	2870	61.1	2617	9	US-10-982-512-4	Sequence 4, Appl
24	2870	61.1	3716	9	US-10-433-757-30	Sequence 30, Appl
25	2870	61.1	4219	3	US-09-976-674-28	Sequence 28, Appl
26	2870	61.1	4219	9	US-10-982-512-28	Sequence 28, Appl
27	2870	61.1	4302	3	US-09-976-674-24	Sequence 24, Appl
28	2870	61.1	4302	9	US-10-982-512-24	Sequence 24, Appl
29	2863	60.9	3000	7	US-10-415-122-1	Sequence 1, Appl
30	2862	60.9	2660	7	US-10-072-012-225	Sequence 225, App
31	2862	60.9	2660	7	US-10-072-012-225	Sequence 225, App
32	2835	60.3	2495	7	US-10-415-122-8	Sequence 8, Appl
33	2833	60.3	3287	7	US-10-415-122-3	Sequence 3, Appl
34	2820.5	60.0	4180	3	US-09-976-674-36	Sequence 36, Appl
35	2820.5	60.0	4180	9	US-10-982-512-36	Sequence 36, Appl
36	2820.5	60.0	4263	3	US-09-976-674-34	Sequence 34, Appl
37	2820.5	60.0	4263	9	US-10-982-512-34	Sequence 34, Appl
38	2649	56.4	4076	3	US-09-976-674-32	Sequence 32, Appl
39	2649	56.4	4076	9	US-10-982-512-32	Sequence 32, Appl
40	2649	56.4	4159	3	US-09-976-674-30	Sequence 30, Appl
41	2649	56.4	4159	9	US-10-982-512-30	Sequence 30, Appl
42	2638	55.1	2801	5	US-10-098-841-100	Sequence 100, App
43	2599.5	55.3	4037	3	US-09-976-674-40	Sequence 40, Appl
44	2599.5	55.3	4037	9	US-10-982-512-40	Sequence 40, Appl
45	2599.5	55.3	4120	3	US-09-976-674-38	Sequence 38, Appl

ALIGNMENTS

RESULT 1
US-10-054-776-1
; Sequence 1, Application US/10054776
; Publication No. US20030165818A1
; GENERAL INFORMATION:
; APPLICANT: Mark Robert Edbrooke
; APPLICANT: Alan Peter Lewis
; TITLE OF INVENTION: NOVEL PROTEIN
; FILE REFERENCE: QG1042US
; CURRENT APPLICATION NUMBER: US/10/054,776
; CURRENT FILING DATE: 2002-01-23
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 2649
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2649)
US-10-054-776-1

Alignment Scores:
Pred. No.: 0
Score: 4700.00
Percent Similarity: 100.0%
Best Local Similarity: 100.0%
Query Match: 6
DB: 6
Length: 2649
Matches: 882
Conservative: 0
Mismatches: 0
Indels: 0
Gaps: 0

US-10-825-632-1 (1-882) x US-10-054-776-1 (1-2649)

Qy 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
Db 1 ATGGCAGCAGCAATGGAAACAGAAACAGCTGGGTGGTGGATATTTGAAATTCGGGACTGT 60
Qy 21 GluGluAenIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
Db 61 GAGGAGAAATATGAATCACAGATCGGCCCTAAATGGAGCCCTTTATATGTTGAGCGGTAT 120
Qy 41 SerTrpSerGlnLeuLysLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
Db 121 TCCTGGAGTCAGCTTAAAGAGCTGTTCGATACCAAGAAATATATCGGCTACATGATG 180
Qy 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
Db 181 GCTAAGGCACACATGATATTCATGTTTGTGAAGAGAAATGATCCAGATGGACCTCATTC 240
Qy 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
Db 241 GACAGAAATCTATTACCTTGGCATGTCTGTGTGAGAACAGAGAAATATACACTGTTTTATTC 300
Qy 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
Db 301 GAAATTCCTCAAAATATCAATAGACAGCAGTCTTAATGCTCTCTTGGAGCCCTCTTTTG 360
Qy 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg 140
Db 361 GATCTTTTTCAGGCAACACTGGACTATGGAAATGTAATCTCGAGAAAGAAACTATTAA 420
Qy 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
Db 421 GAAAGAAACCGCATTTGGAACAGTCGGNAATGCTCTTACGATATATCAACAGGAAGTGA 480
Qy 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
Db 481 ACATTTCTGTTTCAAGCCGGTAGTGGNAATTTACGTAAGAATGGAGGCCACAGGA 540
Qy 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
Db 541 TTTAGCCAAACACTTTAAGCCCAATCTAGTGGAACTAGTGTGCTCCCAACATACGGATG 600
Qy 201 AspProLysLeuCysProAlaAspProAspTyrIleAlaPheIleHisSerAsnAspIle 220
Db 601 GATCCAAATTTATGCCCTGTGATCCAGACTGGATTTGCTTTTATACATAGCAACGATAT 660
Qy 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu 240
Db 661 TGGATATCTAAATCGTAACAGAGAAAGAGAGACTCACTTATGTGCAACATGAGCTA 720
Qy 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu 260
Db 721 GCCAACATGGGAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAAGAAA 780
Qy 261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280
Db 781 TTTGATAGATATTCTGGCTATTGGTGGTGTCCAAAAGCTGAAACAACTCCCAAGTGGTGGT 840
Qy 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300
Db 841 AAAATTTCTAGAAATCTATATGAAGAAATGATGAATCTGAGGTGGAAATATTATTCATGTT 900
Qy 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320
Db 901 ACATCCCTATATTTGGAAACAGAGAGGCAGATTCAATTCGGTTATCTCTTAAACAGGTACA 960
Qy 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
Db 961 GCAATCTCTAAAGTCATCTTTTAAAGATGTCAGAAATTAATGATGATGCTGGAAGAGGATC 1020
Qy 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
Db 1021 ATAGATGTCATAGATAAGGAACCTAATTCACCTTTTGGATTTCTATTGGAAGAGTTGAA 1080

Qy 361 TyrIleAlaArgAlaGlyTyrThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
Db 1081 TATATTGCCAGAGCTGGATGGACTCTCTGAGGGAATATATGCTTGGTCCATCTACTAGAT 1140
Qy 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
Db 1141 CGCTCCAGACTCGCTTCGAGATAGTGTGATCTCACTGAATTAATTAATCCAGTAGAA 1200
Qy 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
Db 1201 GATGATGTTATGGAAAGCAGAGACTCAATGAGTCAGTGCCTGATTTCTGTGAGCCACTA 1260
Qy 421 IleIleTyrGluGluThrThrAspIleTyrIleAsnIleHisAspIlePheHisValPhe 440
Db 1261 ATTATCTATGAAGAAACACAGACTCTGGATAAATATCCATGACATCTTTCATGTTTT 1320
Qy 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
Db 1321 CCCCAAGTCAAGAGAGGAAATTTAGTTTATTTTTCCTCTGAAATGCAAAAACAGGTTTC 1380
Qy 461 ArgHisLeuTyrLysIleThrSerIleLeuLeuGluSerLysTyrLysArgSerSerGly 480
Db 1381 CGTCATTTTATACAAATTTACATCTATTTTAAAGAAAGCAAAATATATAACGATCCAGTGT 1440
Qy 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 500
Db 1441 GGGCTGCTGCTCCCAAGTGATTTCAAGTGTCTTATCAAGAGAGAGATAGCAATTTACAGT 1500
Qy 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
Db 1501 GGTGAATCGGAAGTTCTTTGGCCGCATGGATCTAATATCCAAGTTGATGAAGTCAGAAG 1560
Qy 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisIleLeuTyrValValSer 540
Db 1561 CTGGTATATTTTGAAGCCACCAAGACTCCCTTTAGAGCATCACCTGTACGTAGTCACT 1620
Qy 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
Db 1621 TACGTAATCTCGAGAGGTGACAAAGCTGACTGACCGTGGCTACTCACATTTCTTGCTGC 1680
Qy 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
Db 1681 ATCAGTCAGCACTGTGACTCTTTTAAAGTAAGTATAGTAACCAAGAAAGATCCCACTGT 1740
Qy 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
Db 1741 GTGTCCCTTTTAAAGCTATCAAGTCTTGAAGATGACCAACTTGCAAAACAAGGAATTT 1800
Qy 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 620
Db 1801 TGGGCCACCATTTTGGATTCAGCAGGTCCTCTTCTGACTATACTCTCCAGAAATTTTC 1860
Qy 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
Db 1861 TCTTTTGAAGTACTACTGGATTTTACATTTGATGGGATGCTCTACAGGCTCATGATCTA 1920
Qy 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 660
Db 1921 CAGCTCGAAAGAAATATCTTACTGTCTGTTCATATATGTTGGTCTCTCAGGTGCAAGTTG 1980
Qy 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680
Db 1981 GTGAATATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTAGCCTCTCTAGGT 2040
Qy 681 TyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 700
Db 2041 TATGTGGTTGTAGTAGATAGACCAACAGGGGATCTGTCCAGGAGGCTTAAATTTGAAGC 2100
Qy 701 AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr 720
Db 2101 GCTTTTAAATATAAATGGGTCAANTAGAAATGACGATCAGGTGGAAGGACTCCAAATAT 2160
Qy 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyr 740

Db 2161 CTAGCTTCGATATGATTTCAATGACTTGTAGATCGTGGGCATCCAGGCTGGTCTAT 2220
Qy 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
Db 2221 GGAGGATACCTCTCCCTGATGCATTAATGCAGAGGTACAGATATCTTCAGGGTGGCTATT 2280
Qy 761 AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780
Db 2281 GCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACAGGAACGTTATATG 2340
Qy 781 GlyHisProAspGlnAsnGlnGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
Db 2341 GGTCAACCCCTGACAGAAATGAACAGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAA 2400
Qy 801 LysPheProSerGlnProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
Db 2401 AAGTTCCCTCTGACCAATCGTTTACTGTCTTACATGTTTCTTGGATGAGATGTC 2460
Qy 821 HisPheAlaHisThrSerIleLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
Db 2461 CATTTTGCACATACCATGATATATTACTGAGTTTCTTGTAGGGCTGGAAGCCATATGAT 2520
Qy 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
Db 2521 TTACAGATCTATCTCTCAGAGAGACACAGCATTAAGAGTTCCTGAATCTGGAGAACATTAT 2580
Qy 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeuLys 880
Db 2581 GAATGCATCTTTTGTGCACTACCTTCAAGAAACCTTGATGATCAGCTATGCTGCTTAAAA 2640
Qy 881 ValIle 882
Db 2641 GTGATA 2646

RESULT 2

US-10-170-789-39
; Sequence 39, Application US/10170789
; Publication No. US20030180930A1
; GENERAL INFORMATION:
; APPLICANT: Rachel E. Meyers
; APPLICANT: Olandt, Peter J.
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Williamson, Mark
; APPLICANT: Weich, Nadine
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASE, PHOSPHATASE,
; FILE REFERENCE: 10448-191001
; CURRENT APPLICATION NUMBER: US/10/170,789
; CURRENT FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: US 09/797,039
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: PCT/US01/06525
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/186,061
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 09/882,166
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: PCT/US01/19269
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/212,078
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/934,406
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: PCT/US01/26052
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/226,740
; PRIOR FILING DATE: 2000-08-21
; PRIOR APPLICATION NUMBER: US 09/861,801
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: PCT/US01/16549
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: US 60/205,508
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 09/801,267
; PRIOR FILING DATE: 2001-03-06
; PRIOR APPLICATION NUMBER: PCT/US01/07138
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: US 60/187,454
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: US 09/829,671
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: PCT/US01/40483
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: US 60/197,508
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: US 09/961,721
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: PCT/US01/29904
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: US 60/235,023
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US 10/045,367
; PRIOR FILING DATE: 2001-11-07
; PRIOR APPLICATION NUMBER: US 60/246,561
; PRIOR FILING DATE: 2000-11-07
; PRIOR APPLICATION NUMBER: US 09/801,275
; PRIOR FILING DATE: 2001-03-06
; PRIOR APPLICATION NUMBER: PCT/US01/07074
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: US 60/187,420
; PRIOR FILING DATE: 2000-03-07
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 39
; LENGTH: 2649
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-170-789-39
Alignment Scores:
Pred. No.: 0 Length: 2649
Score: 4700.00 Matches: 882
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 6 Gaps: 0
US-10-825-632-1 (1-882) x US-10-170-789-39 (1-2649)
Qy 1 MetAlaAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
Db 1 ATGGCAGCAGCAATGGAAACAGACAGCTGGGTGTGAGATATTTGAAACTGCGGACTGT 60
Qy 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
Db 61 GAGGAGAATATTGAATCACAGGATCGGCTAAATTTGGAGCCCTTTTATGTGTGACGGGTAT 120
Qy 41 SerTyrSerGlnLeuLysLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
Db 121 TCTGTGAGTCAGCTTTAAAAAGCTGTTCGCCATACCAAGAAATATCATGGCTACATATG 180
Qy 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
Db 181 GCTAAGGCACCATCATGATTTTCATGTTTGTGAAGGAATGATCCAGATGGACCTCATTTCA 240
Qy 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
Db 241 GACAGAATCTATTACCTTCCCATGTCTGTGAGAACACAGAGAAATATACACTGTTTATTTCT 300
Qy 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
Db 301 GAAATTCCTCCAAACTATCAATGAGCAGGATCTTAATGCTCTCTTGGAGCCCTTTTITG 360
Qy 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140

Db	361	GATCTTTTTCAGGCAACACTGGACTATGGAAATGATTTCTCGAGAGAGAAAGAACTATTAAAGA	420	Db	1441	GGGCTGCCTCTCCAAAGTGATTTTCAAGTGTCTCTATCAAAAGAGGAGATAGCAATTACCAGT	1500
Qy	141	GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly	160	Qy	501	GlyGlnTyrGlnValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg	520
Db	421	GAAGAAGAAACGATTTGGAACAGTCGGAATTCCTTCTACGATATATCAACAAGGAAGTGA	480	Db	1501	GGTGAATGGGAAGTCTTGGCCGGCATGGATCTAATATCAAGTTGATGAAGTCAGAAGG	1560
Qy	161	ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValValAspGlyGlyProGlnGly	180	Qy	521	LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisGlyLeuTyrValValSer	540
Db	481	ACATTTCTGTTTCAAGCCGGTAGTGGAAATTTATCATCGTAAAGATGAGAGGCCCAAGGA	540	Db	1561	CTGGTATATTTTGAAGCCACCAAGACTCCCTTTTAGAGCATCACTCTGCTAGTACGT	1620
Qy	181	PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet	200	Qy	541	TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys	560
Db	541	TTTAGCCAAACCTTTTAAGCCCACTCTAGTGGAACTAGTTGTCCCAACATACGGATG	600	Db	1621	TACGTAAATCTCGAGAGGTGACAGCGTACGACCGTGGCTACTACATCTCTTGCTGC	1680
Qy	201	AspProLysLeuCysProAlaAspProAspTyrIleAlaPheIleHisSerAsnAspIle	220	Qy	561	IleSerGlnHisCysAspPheIleSerLysTyrSerAsnGlnLysAsnProHisCys	580
Db	601	GATCCAAATATATGCCCTGCTGATCCAGACTGGATTCCTTTATACATAGCAACGATAT	660	Db	1681	ATCAGTCAGCACTGTGACTCTCTTTATAGTAAGTAACTAGTAACACAGAAATCCACACTGT	1740
Qy	221	TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu	240	Qy	581	ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe	600
Db	661	TGGATATCTAAATCGTAACAGAGAGAAAGAGACTCACATTATGTGCACANTGAGCTA	720	Db	1741	GTGTCCCTTTACAAAGCTATCAAGTCTCGAAGATGACCCAACTTGCACAAACAAGGAATTT	1800
Qy	241	AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu	260	Qy	601	TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe	620
Db	721	GCCAAACATGGGAAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTCTCCAGAGAA	780	Db	1801	TGGGCCACCAATTTGGATTTCAGCAGGTCTCTCTCTGACTATATCTCTCCAGAAATTTTC	1860
Qy	261	PheAspArgTyrSerGlyTyrTyrTyrCysProLysAlaGluThrThrProSerGlyGly	280	Qy	621	SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu	640
Db	781	TTTGATAGATATCTCGGCTATTTGGTGTGTCCAAAAGCTGAAACAACTCCCACTGGTGT	840	Db	1861	TCTTTTGAAGACTACTACTGGATTTCACATTGTATGGATGCTCTACAAAGCTCATGATCTA	1920
Qy	281	LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal	300	Qy	641	GlnProGlyLysValLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu	660
Db	841	AAAAATCTTAGAAATCTATATGAAGAAATATGATGATCTGAGTGGAATATTTATGTT	900	Db	1921	CAGCCTGGNAAGAAATATCTTACTGTCTGTTCATATATATGCTGCTCTCAGGTGCAAGTTG	1980
Qy	301	ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr	320	Qy	661	ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly	680
Db	901	ACATCCCTATGTTGGAACCAAGAGGCGCAGATTCAATCCGTTATCTATCAACAGGTACA	960	Db	1981	GTGAATATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCCCTAGCTCTCTAGGT	2040
Qy	321	AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle	340	Qy	681	TyrValValValValIleAspAsnArgGlySerCysHisArgGlyLysLeuLysPheGlyGly	700
Db	961	GCAATCTCTAAAGTCACTTTAAAGATGTGCAGAAATATGATGATGCTGAAGGAAGATC	1020	Db	2041	TATGTGGTGTAGTGTAGACAAACAGGGGATCTGTCTCACCCGAGGGCTTAAATTTGAAGGC	2100
Qy	341	IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu	360	Qy	701	AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr	720
Db	1021	ATAGATCTCATAGATAAGGAACATAATTCACCTTTTCAGATTTCTATTTGAAGGAGTTGAA	1080	Db	2101	GCCTTTAAATATAAAATGGGTCAATAGAAATTCAGCGATCAGGTGGAAGGACTCCAATAT	2160
Qy	361	TyrIleAlaArgAlaGlyTyrThrProGluGlyLysTyrAlaTyrSerIleLeuLeuAsp	380	Qy	721	LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyr	740
Db	1081	TATATTGCCAGAGCTGGATGGACTCTGAGCGGAAATATGCTTGGTCCATCTCTACTAGAT	1140	Db	2161	CTAGCTTCTCGATATGATTTTCAATGACTTAGATCTGTGGGCATCCACGCTGGTCTAT	2220
Qy	381	ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu	400	Qy	741	GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle	760
Db	1141	CGCTCCAGACTCGCTCGAGATAGTGTGTGATCTCCTGATGATTTATTTATCCAGTAGNA	1200	Db	2221	GGAGGATACCTCTCCCTGATGGCATTAATGAGAGGTGAGATATCTTCAGGGTGTCTTT	2280
Qy	401	AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu	420	Qy	761	AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet	780
Db	1201	GATGATGTTATGGAAGGAGAGACTCATGTAGTCAGTGGCTGATTTCTGTGAGGCCACTA	1260	Db	2281	GCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACAGGAACGTTATATG	2340
Qy	421	IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe	440	Qy	781	GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu	800
Db	1261	ATTATCTATGAAGAAACAACAGACATCGGATGAAATATCCATGACATCTTTCATGTITTT	1320	Db	2341	GGTCACCTGACCAAGTAATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAA	2400
Qy	441	ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe	460	Qy	801	LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal	820
Db	1321	CCCCAAAGTCACGAAGAGGAAATGAGTTTATTTTGGCTCTGATGCAATGCAAAACAGGTTTC	1380	Db	2401	AAGTTCCCTCTCAACCAAAATCGTTTACTGCTCTTACATGGTTCCTCGATGAGAATGTC	2460
Qy	461	ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly	480	Qy	821	HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp	840
Db	1381	CGTCATTTATACAAAATTAATCTATATTTTAAAGGAAGCAATATATAACCGATCCAGTGGT	1440	Db	2461	CATTTTGCACATACCATGATATTTACTGAGTTTTTTTAGTGAGGGCTGGAAAGCCATATGAT	2520
Qy	481	GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer	500	Db	2521	TTACAGATCTATCTCTCAGGAGAGACACACAGATTAAGATTTCTCTGAATCGGAGAACATTTAT	2580

Qy 861 GluLeuHisLeuLeuHisTyrLeuGlnGluLeuLeuGlySerArgIleAlaAlaLeuLys 880
Db 2581 GAACTGTCATCTTTGCACTACCTTCAAGAAAACCTTGGATCAGTATTTGCTGCTCAAAA 2640

Qy 881 ValIle 882
Db 2641 GTGATA 2646

RESULT 3
US-09-976-674-2
; Sequence 2, Application US/09976674
; Patent No. US20020115843A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinaanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 2671
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-976-674-2

Alignment Scores:
Pred. No.: 0 Length: 2671
Score: 4700.00 Matches: 882
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 3 Gaps: 0

US-10-825-632-1 (1-882) x US-09-976-674-2 (1-2671)

Qy 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
Db 8 ATGGCAGCAATGGAACAGACAGCTGGGTGGATGATTTGAACTCGGACTGT 67

Qy 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
Db 68 GAGGAGAATATTGAATCACAGGATCGCCCTAAATTTGGAGCCCTTTTATGTTGACGGTAT 127

Qy 41 SerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
Db 128 TCCTGGAGTCAGCTTAAAGCTGCTTGGCGATACAGAAATATCATGGCTACATGATG 187

Qy 61 AlalysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
Db 188 GCTAAGCACCACATGATTTTCATGTTTGGAGAGGATGATCCAGATGGACCTCATTTCA 247

Qy 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
Db 248 GACAGAATCTATTACCTTGCCATGTCGTGGTGAACAGAGAAAATATCATCTGTTTATTTCT 307

Qy 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
Db 308 GAAATTCCTCAAACTATCAATAGACAGCAGCTTCTTAAATGCTCTCTTGGAGCCCTTTTGT 367

Qy 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg 140
Db 368 GATCTTTTTCAGGCACACATGGAATGATTTCTCGAGAGAGAACTATTAGA 427

Qy 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
Db 428 GAAAGAAAACGCATTTGGAAACAGTCGGAATTTGCTTCTTACGATTATCACCAAGGAAGTGA 487

Qy 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
Db 488 ACATTTCTGTTTCAAGCCGGTAGTGAATTTATACGCTAAAGAGATGGAGGGCCCAAGGA 547

Qy 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
Db 548 TTTACGCACACACTTTTAAGGCCCAATCTAGTGGAACTAGTTGTGCCAACATACGATG 607

Qy 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
Db 608 GATCCAAAATTTATGCCCTGCTGATCCAGACTGGATTGCTTTATACATAGCAACGATATT 667

Qy 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu 240
Db 668 TGGATATCTAACATCTGTAACACAGAGAAGAGAGACTCCTTATGTGCACAAATGAGCTA 727

Qy 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu 260
Db 728 GCCAACATGGAAGAAGATGCCAGATCAGCTGGAGTCGCTACTTTGTTCTCCAAGAGAA 787

Qy 261 PheAspArgTyrSerGlyTyrTrpCysProLysAlaGluThrThrProSerGlyGly 280
Db 788 TTTGATAGATATTCTGCTATTGCTGCTGCTCCAAAAGCTGAAACAACTCCCAAGTGGTGT 847

Qy 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300
Db 848 AAAATCTTGAATTTCTATATGAAGAAAATGATGATCTGAGGTGGAAATTTATTCATGTT 907

Qy 301 ThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThr 320
Db 908 ACATCCCTATGTTGGAACAAAGGAGGAGGAGATTCAITCCGTTATCTTAAACACAGGTACA 967

Qy 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
Db 968 GCAATCTCTAAGTCACTTTTAAAGATGTCAGAAATAATGATGCTGAAGGAGGATC 1027

Qy 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
Db 1028 ATAGATGTCATAGATAAGGAACCTAATTCACCTTTTGAGATTCTATTTTGAAGGAGTTGAA 1087

Qy 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
Db 1088 TATATTGCCAGACTGGATGGACTCTCTGAGGAAAATATGCTGGTCTCATCTACTAGAT 1147

Qy 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
Db 1148 CGCTCCCAAGACTCGCTGACAGATAGTGTGATCTCACCTGAATTTATTTATCCAGTAGAA 1207

Qy 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
Db 1208 GATGATGTTATGGAAGGAGAGACTCATTTGAGTCAGTGCCTGATTCTGTGACGCCACTA 1267

Qy 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
Db 1268 ATTATCTATCAAGAAAACACAGACATCTCGATAAATATCCATGACATCTTTTCATGTTTTT 1327

Qy 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
Db 1328 CCCCAGAGTCACCAAGAGGAAATTTGAGTTTATTTTGGCTCTGAATGCAAAAACAGGTTTC 1387

Qy 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
Db 1388 CGTCATTTATACAAAATTACATCTATTTAAAGGAAAGCAAAATATAAACGATCCAGTGGT 1447

Qy 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 500
Db 1448 GGGCTGCTGCTCCTCAAGTGAATTTCAAGTGTCTCTATCAAGAGGAGAGATAGCAATACCAGT 1507

Qy 501 GlyGluThrProGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
Db 1508 GGTGAATGGGAAGTTCTTGGCCGGCATGATCTAATATATCAAGTTGATGAAGTCAAGG 1567

521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer 540
 1568 CTGGTATATTTTGAAGGACCAAGAGCTCCCTTTAGAGCATCACTGTACGTAGTCAGT 1627
 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
 1628 TACGTAAATCTCGAGAGGTGACAGGCTGACCTGACCTGGCTACTCATCTTCTGTGTC 1687
 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
 1688 ATCAGTCAGCACTGTGACTCTTTTATAAGTAAGTATAGTAACCAAGAAATCCCACTGT 1747
 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
 1748 GTGTCCTTTTCAAGACTATCACTCTCTGAAGTATGAGTCCCAAGCAAGAAATTT 1807
 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluLysPhe 620
 1808 TGGGCCACATTTTGGATTGACAGGTCCTCTCTGACTACTCTCCAGAAATTTTC 1867
 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
 1868 TCTTTTGAAGTACTTGGATTTTACATTTGATATGGATGCTCTACAAGCTCATGATCTA 1927
 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGluValGlnLeu 660
 1928 CAGCTTGAAGAAATATCTTACTGTCTGTATATATGTTGGTCTCTCAGGTGAGTTG 1987
 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680
 1988 GTGAATAATTCGATTTTAAAGGAGTCAAGTATTTTCGCTTGAATACCTACCTCTCTAGGT 2047
 681 TyrValValValIleAspAsnArgGlySerCysHisArgGlyLysPheGluGly 700
 2048 TATGTGTTGTAGTGATAGACAACAGGGGATCTCTGTCACCGAGGGCTTAAATTTGAAGGC 2107
 701 AlaPheLysTyrLysMetGlyGlnIleGluLysAspGlnValGluGlyLeuGlnTyr 720
 2108 GCCTTTATATATAAATGGGTCAATAGAAATGACGATCAGGTGGAAGGACTCCCAATAT 2167
 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrPheSerTyr 740
 2168 CTAGCTTCTCGATATGATTTTCAATGACTTGTGAGTCTGAGGATCCACGGCTGCTCTAT 2227
 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
 2228 GGAGGATACCTCTCCCTGATGGCATTAATGCAGAGGTGAGATATCTTCAGGGTTGCTATT 2287
 761 AlaGlyAlaProValThrLeuThrPheTyrAspThrGlyTyrThrGluArgTyrMet 780
 2288 GCTGGGGCCCCAGTCACTCTGGATCTTCTATGATACAGGATACAGGAACTTATATG 2347
 781 GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
 2348 GGTCACTTACCAAGAAATGAACAGGGCTATCTAGGATCTGAGGATCTGGCCATGCAAGCAAA 2407
 801 LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
 2408 AAGTTCCCTCTGAACCAATCGTTTACTGCTCTTACATGTTTTCCTGGATGAGATGTC 2467
 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
 2468 CATTTTGCACATACCAAGTATTTACTGAGTTTTCCTGAGGGCTGGAAGCCATATGAT 2527
 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
 2528 TTACAGATCTATCTCAGGAGAGACACAGCAAGAGTTCCTGATCGGGAGAACATTTAT 2587
 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLys 880
 2588 GAATGTCATCTTTTGGCACTACCTTCAAGAAACCTTGGATCACGTAATGCTGCTCTAAAA 2647
 881 ValIle 882

Db 2648 GTGATA 2653
 RESULT 4
 US-10-982-512-2
 ; Sequence 2, Application US/10982512
 ; Publication No. US2005009081A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Q1, Steve
 ; APPLICANT: Akinsanya, Karen
 ; APPLICANT: Riviere, Pierre
 ; APPLICANT: Junien, Jean-Louis
 ; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
 ; FILE REFERENCE: 70669
 ; CURRENT APPLICATION NUMBER: US/10/982,512
 ; CURRENT FILING DATE: 2004-11-05
 ; PRIOR APPLICATION NUMBER: US/09/976,674
 ; PRIOR FILING DATE: 2001-10-12
 ; PRIOR APPLICATION NUMBER: US 60/240,117
 ; PRIOR FILING DATE: 2000-10-12
 ; NUMBER OF SEQ ID NOS: 61
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 2
 ; LENGTH: 2671
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-982-512-2
 Alignment Scores:
 Prod. No.: 0 Length: 2671
 Score: 4700.00 Matches: 882
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 9 Gaps: 0
 US-10-825-632-1 (1-882) x US-10-982-512-2 (1-2671)
 Qy 1 MetAlaAlaMetGluThrGluLeuGlyValGluIlePheGluThrAlaAspCys 20
 Db 8 ATGGCAGCAGCATATGGAACACAGACGCTGGTGTGAGATATTTGAACTCGGACTGT 67
 Qy 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
 Db 68 GAGGAGAAATATTGAATACAGATCGGCTTAATTTGAGGCTTTTATGTTGAGCGGTAT 127
 Qy 41 SerTyrSerGlnLeuLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
 Db 128 TCTGGAGTCAGCTTAAAGAGCTGCTGCCATACCAAAATATATCATGCTACATGATG 187
 Qy 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
 Db 188 GCTAAGGCCACCATGATTTTCATGTTGTGAAGAGGAATGATCCAGATGGACCTCATCA 247
 Qy 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
 Db 248 GACAGAAATCTATTACCTTGCCATGTCGTGAGAACAGAGAAATACTGTTTATTTCT 307
 Qy 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTyrLysProLeuLeu 120
 Db 308 GAATTTCCCAAAATATCAATAGACAGCAGCTTAATGCTCTCTTGGAGCCCTCTTTTG 367
 Qy 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg 140
 Db 368 GATCTTTTTCAGGCAACACTGGACATATGGAATGATTTCTCGAGAAGAAGAACTATTAGA 427
 Qy 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
 Db 428 GAAGAAGAACCATTTGGACAGCTCGGAATGCTTCTTACGATTATCACCAAGGAAGTGA 487
 Qy 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
 Db 488 ACATTTCTGTTTTCAGCCGGTAGTGGATTTTATCAGCTAANAAGATGGAGGGCCCAAGGA 547

Qy 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
Db 548 TTTAGCGCAACACCTTTAAGGCCCAATCTAGTGGAAACTAGTTGTCCCAACATACGGATG 607
Qy 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
Db 608 GATCCAAAATTATGCGCTGCTGATCCAGACTGGATTGCTTTATACATGCAACAGATATT 667
Qy 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTrpValHisAsnGluLeu 240
Db 668 TGGATATCTTAACATCGTAACCGAGAGAGAGACTCACTTATGTCACATGAGCTA 727
Qy 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu 260
Db 728 GCCAACATCGAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAAGAA 787
Qy 261 PheAspArgTrpSerGlyTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280
Db 788 TTTGATAGATATTTCTGGCTATTGGTGGTGTCCAAAGCTGAAACAACTCCCACTGGTGT 847
Qy 281 LysIleLeuArgIleLeuTrpGluGluAsnAspGluSerGluValGluIleHisVal 300
Db 848 AAAATTCCTTAGAATTCATATGAGAGAAATGATGAATCTGAGTGGGAAATATTATCATGTT 907
Qy 301 ThrSerProMetLeuGluThrArgAlaAspSerPheArgTrpProLysThrGlyThr 320
Db 908 ACATCCCTTATGTTGGAAACAGAGGCGCAGATTCATTCCTGTTATCTCTAAACAGGTACA 967
Qy 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
Db 968 GCAATCTCTAAAGTCACITTTAAGATGTGAGAAATATGATGATGCTGAAGGAAGGATC 1027
Qy 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluValGlu 360
Db 1028 ATAGATGTCATAGAAAGCACTAATTCACCTTTTGAGATTCATTTTGAAGGAGTTGAA 1087
Qy 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTrpAlaTrpSerIleLeuLeuAsp 380
Db 1088 TATATTGCCAGAGCTGGATGGACTCTGAGGGAANAATATGCTGGTCCATCTCTAGAT 1147
Qy 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
Db 1148 CGCTCCACAGCTCGCTCGCAGATAGTGTGTGATCTCACCTGAATTTATTTATCCAGTAGAA 1207
Qy 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
Db 1208 GATGATGTTATGGAAGGCGAGAGACTCAATTGAGTCAGTCGCTGATTCGTGAGCGCCACTA 1267
Qy 421 IleIleThrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
Db 1268 ATTATCTATGAGAGAAACAACAGACATCTGGATTAATATCCATGACATCTTTCATGTTTT 1327
Qy 441 ProGlnSerHisGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
Db 1328 CCCCAAGTCACGAGAGGAATTCAGTTTATTTTGGCTCTGTAATGCAAAAACAGGTTTC 1387
Qy 461 ArgHisLeuTrpLysIleThrSerIleLeuLysGluSerLysTrpLysArgSerSerGly 480
Db 1388 CGTCATTTATACAAAATTACATCTATTTAAAGGAAGCAAAATATAAACGATCCAGTGGT 1447
Qy 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluIleAlaIleThrSer 500
Db 1448 GGCGCTCGCTGCTCCAAAGTGAATTTCAAGTGTCTATCAAAAGGAGATAGCAATTAACAGT 1507
Qy 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
Db 1508 GGTGAATGGGAAGTTCTTGGCCGGCATGGATCTAATATCCAAAGTTGATGAAGTCAGAGG 1567
Qy 521 LeuValTrpPheGluGlyThrLysAspSerProLeuGluHisHisLeuTrpValValSer 540
Db 1568 CTGGTATATTTGAAGGCCCAAGACTCCCTTTAGAGCATCACCTGTAGTACGTAGTCAGT 1627

RESULT 5

Qy 541 TyrValAsnProGlyGluValThrArgIleThrAspArgGlyTyrSerHisSerCysCys 560
Db 1628 TACGTAAATCCCTGGAGAGGTGACAAAGGCTGACTACCGTGGCTACTCACATTTCTTCTGTC 1687
Qy 561 IleSerGlnHisCysAspPhePheIleSerLysTrpSerAsnGlnLysAsnProHisCys 580
Db 1688 ATCAGTCACACATGTGACTTCTTTATAGTAAGTATATGAACACAGAGAATCCACACTGT 1747
Qy 581 ValSerLeuTrpLysLeuSerSerProGluAspAspProThrCysLysLysThrLysGluPhe 600
Db 1748 GTGTCCCTTTACAAGCTATCAAGTCTCAAGATGCCCAACTTGTCAAAAACAAGGAATTT 1807
Qy 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTrpThrProProGluIlePhe 620
Db 1808 TGGGCCACCATTTTGGATTGACAGGCTCTCTCTGACTATATCTCTCCGAAATTTTC 1867
Qy 621 SerPheGluSerThrThrGlyPheThrLeuTrpGlyMetLeuTrpLysProHisAspLeu 640
Db 1868 TCTTTTGAAGTACTACTGGATTATCATTTGATGGGATGCTCTACAAAGCCTCATGATCTA 1927
Qy 641 GlnProGlyLysLysTrpProThrValLeuPheIleTrpGlyGlyProGlnValGlnLeu 660
Db 1928 CAGCCTGGAAAGAAATATCTACTGTCTGTTCATATATATGCTGCTCCTCAGGTGCTAGTTG 1987
Qy 661 ValAsnAsnArgPheLysGlyValLysTrpPheArgLeuAsnThrLeuAlaSerLeuGly 680
Db 1988 GTGAATAATCGATTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTTAGCCTCTCTAGGT 2047
Qy 681 TyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 700
Db 2048 TATGTGGTTGTGATGATAGAACACAGGGGATCCTGTCCACCGGGCTTAAATTTGAAGGC 2107
Qy 701 AlaPheLysTrpLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTrp 720
Db 2108 GCCTTTAATAATAAATGSGTCAANTAGAAATTCAGATCAGGTGGAAGGACTCCAATAT 2167
Qy 721 LeuAlaSerArgTrpAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTrp 740
Db 2168 CTAGCTTCTCGATATGATTTTCTTACTTGTAGATCTGTGGGATCCACGGCTGCTCTAT 2227
Qy 741 GlyGlyTrpLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
Db 2228 GGAGGATACCTCTCCCTGATGGCATTAATGCAAGAGTCAATATCTTTCAGGGTTGCTATT 2287
Qy 761 AlaGlyValaProValThrLeuTrpIlePheTrpAspThrGlyTrpThrGluArgTrpMet 780
Db 2288 GCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACGTATATG 2347
Qy 781 GlyHisProAspGlnAsnGluGlnGlyTrpTrpLeuGlySerValAlaMetGlnAlaGlu 800
Db 2348 GGTCACTCCCTGACCAGAAATGAACAGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAA 2407
Qy 801 LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
Db 2408 AAGTTCCTCTGAAACAAATCGTTTACTGTCTTACATGGTTTCTCGATGAGAATGTC 2467
Qy 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTrpAsp 840
Db 2468 CATTTTGCACATACCACTATATATCTAGTTTTTTGTAGGGCTGGAAAGCCATATGAT 2527
Qy 841 LeuGlnIleTrpProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTrp 860
Db 2528 TTACAGATCTATCTCTCAGGAGACACAGCAATAGAGTTCCTGAATTCGGGAGAACATTTAT 2587
Qy 861 GluLeuHisLeuLeuHisTrpLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLys 880
Db 2588 GAACGTGATCTTTTGTGACTACCTTCAAGAAAACCTTGGATCAAGTATTGCTGCTCTAAA 2647
Qy 881 ValIle 882
Db 2648 GTGATA 2653

```
US-10-311-035-30
; Sequence 30, Application US/10311035
; Publication No. US20040023243A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: YUE, Henry
; APPLICANT: ELLIOTT, Vicki
; APPLICANT: GANDHI, Ameena R.
; APPLICANT: LAL, Preeti
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: TRIBOULEY, Catherine M.
; APPLICANT: DELEGEANE, Angelo M.
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: NGUYEN, Danielle B.
; APPLICANT: LEE, Ernestine A.
; APPLICANT: HAFALIA, April
; APPLICANT: KHAN, Farrah A.
; APPLICANT: CHAWLA, Narinder K.
; APPLICANT: YAO, Monique G.
; APPLICANT: LU, Dyung Aina M.
; APPLICANT: ARVIZU, Chandra S.
; APPLICANT: TANG, Y. Tom
; APPLICANT: WALSH, Roderick T.
; APPLICANT: AZINZAI, Yalda
; APPLICANT: LU, Yan
; APPLICANT: RAMKOMAR, Jayalaximi
; APPLICANT: REDDY, Roopa
; APPLICANT: DAS, Depopriya
; APPLICANT: KEARNEY, Liam
; APPLICANT: KALLICK, Deborah A.
; TITLE OF INVENTION: Proteases
; FILE REFERENCE: PI-0123 PCT
; CURRENT APPLICATION NUMBER: US/10/311.035
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: 60/212,336; 60/213,955; 60/215,396; 60/216,821; 60/218,946
; PRIOR FILING DATE: 2000-06-16; 2000-06-22; 2000-06-29; 2000-07-07; 2000-07-14
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PERL Program
; SEQ ID NO 30
; LENGTH: 3106
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20040023243A1 7160544CB1
US-10-311-035-30

Alignment Scores:
Pred. No.: 0 Length: 3106
Score: 4700.00 Matches: 882
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: Gaps: 0

US-10-825-632-1 (1-882) x US-10-311-035-30 (1-3106)
Qy 1 MetAlaAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 20
Db 203 ATGGCAGCAGCAATGGAACACAGACAGCTGGGTGTGAGATATTGAAACTGGCGACTGT 262
Qy 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
Db 263 GAGGAGATATTGAAATCAAGATCGGCCTAAATGGAGCCCTTTTATGTTGAGCCGGTAT 322
Qy 41 SerTyrSerGlnLeuLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
Db 323 TCCTGGAGTCAGCTTTAAAGCTGCTTCGCATATACAGAAATATATCATGCTACATGATG 382
Qy 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
Db 383 GCTAAGGCCACCATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTC 442
```

81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
443 GACAGATCTATTACCTTGCATGTCTGGTGAGACAGAGAAAATACACTGTTTTATTCT 502

101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
503 GAAATCCCAAACTATCAATAGAGCAGCAGTCTTAATGCTCTCTTTGGAGCCTCTTTG 562

121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140
563 GATCTTTTTCAGGCAACACTGGACTATGGAAATTCCTCAGAGAAAGACTATTAAAG 622

141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
623 GAAAGAAAACGATTTGGAACTCGGAATTCCTTACGATTATCAGAGAAAGACTGGA 682

161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
683 ACATTTCTGTTTCAAGCCGCTAGTGGAAATTTATCAGCTAAAGATGGAGGCCACAGGA 742

181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
743 TTTACGCAACAACCTTTAAGGCCCAATCTAGTGGAAACTAGTTGTGCCAACATACGATG 802

201 AspProLysLeuCysProAlaAspProAspTyrIleAlaPheIleHisSerAsnAspIle 220
803 GATCCAAAATTTATGCCCTGCTGATCCAGACTGGATTGCTTTTATACATAGCAACGATAT 862

221 TyrIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu 240
863 TGGATATCTTACATCGTAACAGAGAAAGAGAGACTCACTTATGTGCAATGAGCTA 922

241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu 260
923 GCCACATCGAAGAGATGCCAGATCAGCTGGAGTCCCTACCTTTGTTCTTCCAAAGAA 982

261 PheAspArgTyrSerGlyTyrTyrTyrCysProLysAlaGlnThrThrProSerGlyGly 280
983 TTTGATAGATATTCGGCTATTTGGGTGTCGCAAGCTGAACCACTCCCATGGTGGT 1042

281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300
1043 AAAATCTTTAGAAATCTATATGAAGAAATGATGAATCTGAGGTGGAAATTTATTCATGT 1102

301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320
1103 ACATCCCTTATGTTGAAACCAAGGAGGCGCAGATTCAATCCGTTATCTCTAAACAGGTACA 1162

321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
1163 GCAATCTCTAAAGTCACTTTTAAGATGTGAGAAATTAATGATGCTGAAGGAAGGATC 1222

341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
1223 ATAGATGTCATAGATAAGGAACTAATTCACCTTTTGAGATTCCTATTGGAAGGAGTTGAA 1282

361 TyrIleAlaArgAlaGlyTyrThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
1283 TATATTGCCAGAGCTGGATGGACTCTCGAGGGAATAATGCTTGCTCCATCTACTAGAT 1342

381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
1343 CGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGANTTATTATATCCCAAGTAGAA 1402

401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
1403 GATGATGTTATGGAAGGCGAGACTCAATGAGTCAGTGCCTGATTTGTCGACGCCACTA 1462

421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
1463 ATTATCTATGAAGAAACCAACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTT 1522

QY 441 ProGlnSerHisGluGluGluIleGluPheAlaSerGluCysGlyThrGlyPhe 460
Db 1523 CCCCAAGTACGAGAGGAAATTGAGTTTATTTTGGCTCTGAAATGCAAAACAGGTTTC 1582
QY 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerGly 480
Db 1583 CGTCATTTATACAAATTTACATCTATTTTAAAGGAAGCAATATAAAGCATCCAGTGGT 1642
QY 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 500
Db 1643 GGGCTGCTGCTCCAGTGATTTCAAGTGCTCTATCAAGAGAGATAGCAATTTACCAGT 1702
QY 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArg 520
Db 1703 GGTGAATGGGAAGTTCTTTGGCCGCGCATGGATCTAATCAAGTTGATGAAGTCAGAAG 1762
QY 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisIleLeuTyrValValSer 540
Db 1763 CTGGTATATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACCCTGTACGTAGTCAGT 1822
QY 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
Db 1823 TACGTAATCTCGAGAGGTGACCAAGCTGACTGACCGTGGCTACTCACTTTCTGCTGCG 1882
QY 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
Db 1883 ATCAGTCAGCACTGTGACTCTTTTATTAAGTAAATAGTAGTAACCAAGAGAAATCCACACTGT 1942
QY 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
Db 1943 GTGTCCTTTTCAAGCTATCAAGTCTCGAAGATGACCAACTTGCAGAAAACAAAGGAATTT 2002
QY 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 620
Db 2003 TGGGCCACCATTTTGGATTACAGAGGTCTCTTCCGACTATACTCTCCAGAAATTTTC 2062
QY 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
Db 2063 TCTTTTGAAGTACTACCTGGATTTTACATTTGATGGGATGCTCTACAGGCTCATGATCTA 2122
QY 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 660
Db 2123 CAGCTCGAAGAAATATCTTACTGTCTGTTCATATATGCTGCTCCTCAGGTGCAATG 2182
QY 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680
Db 2183 GTGAATAATCGGTTTAAAGGAGTCAAGTATTTTCGGCTTGAATACCCCTAGCCTCTTAGT 2242
QY 681 TyrValValValValIleAspAsnArgGlySerCysHisArgGlyLysLysPheGluGly 700
Db 2243 TATGTGGTTGTAGTATAGACACAGGGGATCTGTCACGAGGGCTTAAATTTGAGGC 2302
QY 701 AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr 720
Db 2303 GCCTTTAAATATAAATGGGTCAATAGAAATTCAGCATCAGTGGAGGACTCCCAATAT 2362
QY 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyr 740
Db 2363 CTAGCTTCTCGATATGATTTTCATTTGACTTATGCTGTGGCATCCACGGCTGCTCTAT 2422
QY 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
Db 2423 GGAGGATACCTCTCCCTGATGGCATTAAATGACAGAGGTGAGATATCTTCAGGGTGTCTAT 2482
QY 761 AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780
Db 2483 GCTGGGGCCCCAGTCACTCTGTGGATCTTCTATATACAGATACAGGACGCTATATG 2542
QY 781 GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
Db 2543 GGTCAACCTGACCAAGATGAACAGGGCTATTACTAGGATCTGTGGCATGCAAGCAGAA 2602
QY 801 LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820

Db 2603 AGTTCCCTCTGAACCAATTCGTTTACTGCTCTTACATGGTTTCTCGATGGAATGTC 2662
QY 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
Db 2663 CATTTTGCACATACACAGTATATTTACTGAGTTTTTTAGTGAGGGCTGAAAGCCATATGAT 2722
QY 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
Db 2723 TTACAGATCTATCTCTCAGGAGACACAGCATAGAGTTTCTGAATCGGGAGAACATTAT 2782
QY 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLys 880
Db 2783 GAACGTGCATCTTTTGCACCTTCAAGAAAACCTTGGATCACGTATTTGCTGCTCTAAA 2842
QY 881 ValIle 882
Db 2843 GTGATA 2848
RESULT 6
US-10-415-122-5
; Sequence 5, Application US/10415122
; Publication No. US20040053369A1
; GENERAL INFORMATION:
; APPLICANT: THE UNIVERSITY OF SYDNEY
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES
; FILE REFERENCE: FP15217
; CURRENT APPLICATION NUMBER: US/10/415,122
; CURRENT FILING DATE: 2003-08-07
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 5
; LENGTH: 3120
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-415-122-5
Alignment Scores:
Pred. No.: 0 Length: 3120
Score: 4700.00 Matches: 882
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 7 Gaps: 0
US-10-825-632-1 (1-882) x US-10-415-122-5 (1-3120)
QY 1 MetAlaAlaLysGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
Db 214 ATGGCAGCAGCAATGCAACAGAACAGCAGCTGGGTGTGAGATATTTGAAACTGCGGACTGT 273
QY 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
Db 274 GAGGAGAAATTTGAATTCACAGGATCGGCTTAATTTGAGGCTTTTATGTTGAGCGGTAT 333
QY 41 SerTrpSerGlnLeuLysLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
Db 334 TCTTGGAGTCAGCTTAAAGAGCTGCTCCGATACCAAGAAATATCATGCTACATGATG 393
QY 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
Db 394 GCTAAGGCACCATGATTTTCATGTTGTGAAGAGGAATGATCCAGATGACCTCATTC 453
QY 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAanThrLeuPheTyrSer 100
Db 454 GACAGAAATCTTATTCCTGCAATGCTGGTGGAGAACAGAGAAATATACATCTTTTATCT 513
QY 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
Db 514 GAAATTTCCCAAAACTATCATATAGACAGCAGCTTAAATGCTCTCTTGGAGCCTCTTTTG 573
QY 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140

Db 574 GATCTTTTTCAGGCCAACCTGGACTATGGAAATGATTTCTCGAAGAGAACTATTAGA 633
Qy 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
Db 634 GAAGAAGAACCATTTGGACAGTCGGAAATTCCTCTTACGATATATACCAAGGAAGTGG 693
Qy 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
Db 694 ACATTTCTGTTCGAAGCCGGTAGTGGAAATTTATCAGTAAAGATGGAGGGCCACAGGA 753
Qy 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
Db 754 TTTTACGCAACCAACCTTTAAGCCCAATCTAGTGGAAACCTAGTTGTCCCAACATACGGATG 813
Qy 201 AspProLysLeuCysProAlaAspProAspTyrIleAlaPheIleHisSerAsnAspIle 220
Db 814 GATCCAAATTTATGCCCGCTGATCCAGACTGGATTTGCTTTATATACATAGCAACGATATT 873
Qy 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu 240
Db 874 TGGATATCTAATCATCGTAACAGAGAAAGAGAGACTCACATATATGTGCACAAATGAGCTA 933
Qy 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu 260
Db 934 GCCAACATGGGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAGAAGAA 993
Qy 261 PheAspArgTyrSerGlyTyrTyrTrpCysProLysAlaGluThrThrProSerGlyGly 280
Db 994 TTTGATAGATATCTGGCTATTTGGTGGTGTCCAAAGCTGAAACCACTCCCAAGTGGTGGT 1053
Qy 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHisVal 300
Db 1054 AAAATTTCTTAGAATTTATATAGAAAGAAATGATGAATCTGAGGTGGAATATATCATGTT 1113
Qy 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320
Db 1114 ACATCCCCATGTGTGGAAACAAGAGGGCAGATTCAATCCGTATCTCTTAAACAAGGTACA 1173
Qy 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
Db 1174 GCAAAATCCTAAAGTCACCTTTTAAGATGTCCAGAAATATGATGTGCTGGAAGGAAGATC 1233
Qy 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
Db 1234 ATAGATGTCTAGATAAGGAACCTAAATCAACCTTTTGAGATTTCTATTTGAAGAGTTGAA 1293
Qy 361 TyrIleAlaArgAlaGlyTyrThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
Db 1294 TATATTGCCAGAGCTGGATGGACTCTCTGAGGGAAAAATATGCTTGGTCCATCTCTAGAT 1353
Qy 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
Db 1354 CGCTCCAGACTCGCCTACAGATAGTGTTCATCTCACTGAAATTTATTTATCCAGTAGAA 1413
Qy 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
Db 1414 GATGATGTTATGGAAAGGCAGAGACTCATTTGAGTCAGTGCCTGATTTCTGTGAGCCACTA 1473
Qy 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
Db 1474 ATTATCTATGAAGAAACAACAGACTCTGGATAAATATCATGACACTCTTCATGTTTTT 1533
Qy 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
Db 1534 CCCCNAAGTCAAGAGGAATTTAGTTTATTTTGGCTCTGAAATGCAAAACAGGTTTC 1593
Qy 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
Db 1594 COTCATTTATACAAAATTTACATCTATTTTAAAGGAAGCAAAATATAAACGATCCAGTGGT 1653
Qy 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 500
Db 1654 GGGCTGCCTGCTCCAAAGTGATTTCAAGTGTCTTATCAAGAGGAGATAGCAATTTACCAGT 1713

Qy 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
Db 1714 GGTGAATGGGAAGTTCTTTGGCCGCATGGATCTAATATCCAGTTGATGAAGTCAGAAG 1773
Qy 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer 540
Db 1774 CTGGTATATTTTGAAGGACCAAGACTCCCTTTTAGAGCATCCTGTACGTAGTCAAT 1833
Qy 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
Db 1834 TACGTAATATCTCGAGAGGTGACAAGGCTGACGACCGTGGCTACTCACATTTCTGCTGC 1893
Qy 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
Db 1894 ATCAGTTCAGCACTGTGACTCTTTTATTAAGTAAGTATAGTAACCAAGAAATCCACTGT 1953
Qy 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
Db 1954 GTGTCCCTTTTCAAGCTATCAAGTCTCTGAAGAGTACCCTGCTGCAAAACAAGGAATTT 2013
Qy 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluLysPhe 620
Db 2014 TGGGCCACCATTTTGGATTTCAGCAGGTCTCTCTTCTGACTATATCTCTCAGAAATTTTC 2073
Qy 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
Db 2074 TCTTTTGAAGTACTACTGGATTTTACATTTGATGGGATGCTCTACAAGCTCATGATCTA 2133
Qy 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 660
Db 2134 CAGCTCGAAGAAATATCTTACTGTGCTGTTTATATATGTGTGCTCTCAGGTGCAGTTG 2193
Qy 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAsnLeuGly 680
Db 2194 GTGAATATCCGTTTAAAGGAGTCAGTATTTCCGTTGATACCTTAGCTCTCTAGGT 2253
Qy 681 TyrValValValIleAspAsnArgLysSerCysHisArgGlyLeuLysPheGluGly 700
Db 2254 TATGTGGTTGTAGTGATAGACAACAGGGGATCTCTGTCCAGGGGCTTAAATTTGAAGGC 2313
Qy 701 AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr 720
Db 2314 GCCTTTAAATATAAATGGGTCAAAATAGAAATTTGACGATCAGGTGGAAGGACTCCCAATAT 2373
Qy 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyr 740
Db 2374 CTAGCTTCTCGATATGATTTTCAATTGACTTAGATCGTGTGGGCATCCACGGCTGGTCTAT 2433
Qy 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
Db 2434 GGAGGATACCTCTCCCTGATGGCATTAATGCAGAGGTTCAGATATCTTCAAGGTTCCTATT 2493
Qy 761 AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780
Db 2494 GCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGATACAGGAACGTTATATG 2553
Qy 781 GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
Db 2554 GGTCACTCTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAGACAGA 2613
Qy 801 LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
Db 2614 AAGTTCCCTCTGAACCAAAATCGTTACTCTTACTGCTTCTACATGTTTCTTGGATGAAATGC 2673
Qy 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
Db 2674 CATTTTGCATACACAGTATATTTACTGAGTTTTTTTAGTGGGGCTGGAAGCCATATGAT 2733
Qy 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
Db 2734 TTTACAGATCTATCTCTCAGGAGACACAGCATTAAGAGTTCTCTGAATCGGGAGAACATTAT 2793

Qy 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeuLys 880
Db 2794 GAATGCTATCTTTTGCACCTTCAAGAAACCTTGGATCAGTATTGCTCTAAAA 2853
Qy 881 ValIle 882
Db 2854 GTGATA 2859

RESULT 7

US-10-825-632-2
; Sequence 2, Application US/10825632.
; Publication No. US20040191826A1
; GENERAL INFORMATION:
; APPLICANT: ABBOTT, Catherine Anne
; APPLICANT: GORRELL, Mark Douglas
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES
; FILE REFERENCE: PCSB-100-Div. 1
; CURRENT APPLICATION NUMBER: US/10/825,632
; CURRENT FILING DATE: 2004-04-15
; PRIOR APPLICATION NUMBER: US 10/070,464
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: PCT/AU00/01085
; PRIOR FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: AU PQ5709
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: AU PQ2762
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 3120
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-825-632-2

Alignment Scores:

Pred. No.: 0 Length: 3120
Score: 4700.00 Matches: 882
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 8 Gaps: 0

US-10-825-632-1 (1-882) x US-10-825-632-2 (1-3120)

Qy 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
Db 214 ATGCGCAGCAATGGAAACAGACAGCTGGGTGTGAGATATTGAACTGGCGACTGT 273
Qy 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
Db 274 GAGGAGATATTGAAATCAGACAGATCGGCTAAATTTGGAGCCCTTTTATGTTGAGCGGTAT 333
Qy 41 SerTrpSerGlnLeuLysLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
Db 334 TCCTGGAGTCAGCTTAAAGAGCTGTTCGCGATACAGAAATATCATCGCTACATGATG 393
Qy 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
Db 394 GCTAAGGCACCATGATTTCATGTTGTGAAAGAGGAATGTCAGATGGACCTCATTTCA 453
Qy 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
Db 454 GACAGATCTATTACCTTGGCATGCTGTGGTGAGAACAGAGAAATACATGTTTTATTCT 513
Qy 101 GluIleProLysThrIleAsnArgAlaValLeuMetLeuSerTrpLysProLeuLeu 120
Db 514 GAAATCCCAAACTATCATAGAGCAGCAGCTTAAATGCTCTCTTGGAGCCCTCTTTTG 573
Qy 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140
Db 574 GATCTTTTTCAGGCAACACTGGACTATGGAATGATTCTTCGAGAGAGAACTATTAAAGA 633

Qy 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
Db 634 GAAAGAAAAACGATTTGGAAACAGTCGGAATTCCTCTTACGATTATATCAACGAAGTGA 693
Qy 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
Db 694 ACATTTCTGTTTCNAGCCGCTAGTGGNAATTTATCACTGTAAGAGATGAGGCGCCACA 753
Qy 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
Db 754 TTTTACGCAACAACCTTTAAGGCCCAATCTAGTGGAAACTAGTTGCTCCCAACATACGATG 813
Qy 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
Db 814 GATCCAAAATTTATGCCCGCTGATCCAGATGGATTGCTTTTATACATGACAGCATATT 873
Qy 221 TrpIleSerAsnIleValThrArgGluGluArgArgLeuThrTyrValHisAsnGluLeu 240
Db 874 TGGATATCTAACATCGTAACACAGAGAGAGAGAGACTCACTTATGTGCACATGAGCTA 933
Qy 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu 260
Db 934 GCCAACAATGGAAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTTCTCCAGAGAA 993
Qy 261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280
Db 994 TTTGATAGATATTTCTGGCTATTGGTGGTGTCCAAAGCTGAAACAATCCAGTGGTGGT 1053
Qy 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300
Db 1054 AAAATTTCTAGAAATCTATGAGAAATGATGAACTGAGGTGGAAATTTATTCATGTT 1113
Qy 301 ThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThr 320
Db 1114 ACATCCCTTATGTTGGAAACAAGAGGCGCAGATTCACTCCGTATCTTAAACAAGGTACA 1173
Qy 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
Db 1174 GCAATCTCTAAAGTCACTTTTAAAGATGTGAGAAATATGATGATGATGCTGAGGAGGATC 1233
Qy 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
Db 1234 ATAGATGTCATAGATAAGNACTAATTCACCTTTTGAGATCTTATTTGAGGAGTTGAA 1293
Qy 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
Db 1294 TATATTGCCAGAGCTGGATGGACTCTCGAGGGAATAATGCTTGGTCCATCTACTAGAT 1353
Qy 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
Db 1354 CGCTCCAGACTCGCCTACAGATAGTGTGTGATCTCACCTGAATATTATTTATCCCAAGTAA 1413
Qy 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
Db 1414 GATGATGTTATGGAAGGAGAGACTCATTTGAGTCAGTGCCTGATTTCTGTGACGCCACTA 1473
Qy 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
Db 1474 ATTATCTAGAGAAACAACACACATCTGGATAATATCCATGACATCTTTTCATGTTTTT 1533
Qy 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
Db 1534 CCCCAAGTCAAGAGGAAATTTGAGTTTATTTTGGCTCTGATGATGCAATCTTTTCATGTTTTT 1593
Qy 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerTyrLysArgSerSerGly 480
Db 1594 CGTCATTTATACAAAATTTACATCTATTTTAAAGGAAAGCAATATATAACGATCAGTGGT 1653
Qy 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 500
Db 1654 GGGTGGCTGCTCTCAAGTATTTCAAGTGTCTCTATCAAGAGGAGATAGCAATTTACCAGT 1713
Qy 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520

1714 GGTGAATGGGAAGTTCTTTGGCGGCATGGATCTAATATCCAGTGTGATGAAGG 1773

Db 1714 GGTGAATGGGAAGTTCTTTGGCGGCATGGATCTAATATCCAGTGTGATGAAGG 1773
QY 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHieLysLeuTyrValValser 540
Db 1774 CTGGTATATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCACTGTAGTAGTCAGT 1833
QY 541 TyrValLeuProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
Db 1834 TACGTAAATCTCGAGAGGTGACAGAGCTGACTGACCGTGGCTACTACATCTTCTGTCG 1893
QY 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
Db 1894 ATCAGTCAGCACTGTGACTCTTTTATAGTAGTATAGTAAACAGAGNATCCACACTGT 1953
QY 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
Db 1954 GTGTCCTTTTAAAGCTATCAAGTCCTGAAGATGACCACTGTCGAAACAAAGNATTT 2013
QY 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 620
Db 2014 TGGGCACCATTTTGGATTACAGCAGGTCTCTTCTGACTATACTCTCCAGAAATTTTC 2073
QY 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
Db 2074 TCTTTTGAAGTACTACTGGATTTTACATTTGTATGGATGCTCTACAGCCTCATGATCTA 2133
QY 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 660
Db 2134 CAGCCTGGAAGAAATATCTTACTGTGTCTGTATATATGTTGGTCTCCTCAGGTGCAATG 2193
QY 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680
Db 2194 GTGAATATCCGTTTAAAGGAGTCAAGTATTTTCGGTTGAATACCCCTAGCCTCTCTAGT 2253
QY 681 TyrValValValValIleAspAsnArgLysSerCysHisArgGlyLeuLysPheGluGly 700
Db 2254 TATGTGTTGTAGTATAGACACACAGGGGATCTGTCAACGAGGGCTTAAATTTGAAGGC 2313
QY 701 AlaPheLysTyrLysMetGlyGlnIleGluLeuAspGlnValGluGlyLeuGlnTyr 720
Db 2314 GCCTTTAAATATAAATGGTCAATAGAAATTCAGATCAGTGGAGGACTCCCAATAT 2373
QY 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyr 740
Db 2374 CTAGCTTCTCGATATGATTTTCACTTAGATCGTGGGCATCCACGGCTGGTCTAT 2433
QY 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
Db 2434 GGAGGATACCTCTCCCTGATGGCATTAATGCAGAGGTACAGATATCTTCAGGGTGTCTAT 2493
QY 761 AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780
Db 2494 GCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACCTTATATG 2553
QY 781 GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
Db 2554 GGTCAACCTGACCAGAAATCAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAA 2613
QY 801 LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
Db 2614 AAGTTCCTCTCAACCAATCGTTTACTGTCTTCTACATGTGTTTCTGGATGAGATGTC 2673
QY 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
Db 2674 CATTTTGCATACACAGTATATCTAGTATTTTGTAGGGCTGGAAAGCCATATGAT 2733
QY 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
Db 2734 TTACAGATCTATCTCAGGAGACACACAGATACAGATTCCTGAATCGGAGAACATTAT 2793
QY 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeuLys 880

Db 2794 GAACGTGCATCTTTTGCACTACTCTTCAAGAAACCTTGGATCAGCTATTGCTGCTCTAAAA 2853
QY 881 Vallie 882
Db 2854 GTGATA 2859
RESULT 8
US-10-170-789-37
; Sequence 37, Application US/10170789
; Publication No. US20030180930A1
; GENERAL INFORMATION:
; APPLICANT: Rachel E. Meyers
; APPLICANT: Olandt, Peter J.
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Williamson, Mark
; APPLICANT: Welch, Nadine
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASE, PHOSPHATASE,
; AND PROTEASE FAMILY MEMBERS AND USES THEREOF
; FILE REFERENCE: 10448-191001
; CURRENT APPLICATION NUMBER: US/10/170,789
; CURRENT FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: US 09/797,039
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: PCT/US01/06525
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/186,061
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 09/882,166
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: PCT/US01/19269
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/212,078
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/934,406
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: PCT/US01/26052
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/226,740
; PRIOR FILING DATE: 2000-08-21
; PRIOR APPLICATION NUMBER: US 09/861,801
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: PCT/US01/16549
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: US 60/205,508
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 09/801,267
; PRIOR FILING DATE: 2001-03-06
; PRIOR APPLICATION NUMBER: PCT/US01/07138
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: US 60/187,454
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: US 09/829,671
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: PCT/US01/40483
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: US 60/197,508
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: PCT/US01/29904
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: US 60/235,023
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US 10/045,367
; PRIOR FILING DATE: 2001-11-07
; PRIOR APPLICATION NUMBER: US 60/246,561
; PRIOR FILING DATE: 2000-11-07
; PRIOR APPLICATION NUMBER: US 09/801,275
; PRIOR FILING DATE: 2001-03-06
; PRIOR APPLICATION NUMBER: PCT/US01/07074
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: US 60/187,420

; PRIOR FILING DATE: 2000-03-07
 ; NUMBER OF SEQ ID NOS: 63
 ; SOFTWARE: Fast-Seq for Windows Version 4.0
 ; SEQ ID NO 37

; LENGTH: 3143
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (229) ... (2874)
 US-10-170-789-37

Alignment Scores:
 Pred. No.: 0 Length: 3143
 Score: 4700.00 Matches: 882
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 6 Gaps: 0

US-10-825-632-1 (1-882) x US-10-170-789-37 (1-3143)

Qy 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
 Db 229 ATGGCAGCAGCAATGGAAACAGAACAGCTGGGTGTGAGATATTGAAACCTGGCGACTGT 288
 Qy 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
 Db 289 GAGGAGNATATTGAATCACAGGATCGGCCCTAAATTTGGAGCCCTTTTATGTGAGCGGTAT 348
 Qy 41 SerTrpSerGlnLeuLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
 Db 349 TCCTGGAGTCAGCTTAAAGCTGCTTGGCGATACAGAAATATCATGGCTACATGATG 408
 Qy 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
 Db 409 GCTAAGGCCACCATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTTCA 468
 Qy 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
 Db 469 GACAGAAATCTATTACCTTGGCATGCTGTGTGAGAACAGAGAAATATACACTGTTTATCT 528
 Qy 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
 Db 529 GAATATCCCAAACTATCATAGAGCAGCAGCTTAAATGCTCTCTGGAGCCCTCTTTTG 588
 Qy 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg 140
 Db 589 GATCTTTTTCAGGCAACACTGGACTATGGAATGTATTCTCGAGAGAAGAACTATTAAAGA 648
 Qy 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
 Db 649 GAAAGAAAACCGCATGGACAGTCGGAATTTGCTTCTTACGATTATCACCAAGGAAGTGA 708
 Qy 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
 Db 709 ACATTTCTGTTTACGCCGCTAGTGAATTTATCAGTAAAGATGGAGGCCACAGGA 768
 Qy 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
 Db 769 TTTACGCAACAACTTTAAGGCCAATCTAGTGGAAACTAGTTGTGCCCAACATACGGATG 828
 Qy 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
 Db 829 GATCCAAAATATTATGCCCTGCTGATCCAGACTGGATTGCTTTTATACATAGCAACGATATT 888
 Qy 221 TrpIleSerAsnIleValThrArgGluGluArgGluLeuThrTyrValHisAsnGluLeu 240
 Db 889 TGGATATCTTAACATCGTAACAGAGAGAAAGAGACTCATCTTATGTCCACATGAGCTA 948
 Qy 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu 260
 Db 949 GCCAACATGGAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAAGAGAA 1008

Qy 261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280
 Db 1009 TTTGATAGATATTCTGGCTATTGGTGGTGTCCAAAGCTGAACCACTCCCAAGTGGTGGT 1068
 Qy 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHisVal 300
 Db 1069 AAAATTCCTTAGAATTCATATGAGAAATGATGAAATCTGAGGTGGAAATTTATTCATGTT 1128
 Qy 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320
 Db 1129 ACATCCCTATGTTGGAACAAGGAGGCGAGATTCAATCCGTTATCTCTAAACAGGTACA 1188
 Qy 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
 Db 1189 GCAAAATCCTAAAGTCACTTTTAAGATGTTCAGAAATAATGATGATCTGAGGTGGAAATTTATTCATGTT 1248
 Qy 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
 Db 1249 ATAGATGTCTAGATAAGGAACCTAATCAACCTTTTGAGATCTATTTGAAAGGAGTTGAA 1308
 Qy 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
 Db 1309 TATATGCCAGAGCTGGATGGAATCTCTGAGGAAATATGCTTGGTCCATCTACTAGAT 1368
 Qy 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
 Db 1369 CGCTCCAGACTCGCTCCAGATAGTGTGATCTCACCTGAAATATTTATCCAGTAGAA 1428
 Qy 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
 Db 1429 GATGATGTTATGAAAGGAGGAGACTCATTTGAGTCAGTGCCTGATCTGTGACGCCACTA 1488
 Qy 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
 Db 1489 ATTATCTATGAAAGAACACAGACATCTGGATAATATCCATGACATCTTTTCATGTTTT 1548
 Qy 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
 Db 1549 CCCCNAAGTCACCAAGGAGAAATGAGTTATTTTGTCTCTGAATGCAAAACAGGTTTC 1608
 Qy 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
 Db 1609 CGTCATTTATACAAATTTACATCTATTTTAAAGGAAAGCAATATATAACGATCCAGTGGT 1668
 Qy 481 GlyLeuProAlaProSerAspPheLysCysProLysGluGluIleAlaIleThrSer 500
 Db 1669 GGGCTGCTCTCCAGTGAATTTCAAGTGTCTTATCAAGAGAGATAGCAATTTACCACT 1728
 Qy 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
 Db 1729 GGTGAATGGGAGTTCTTGGCCGCATGATCTAATATCCAAAGTTGATGAAGTCAGAAGG 1788
 Qy 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisIleSerTyrValValSer 540
 Db 1789 CTGTATATTTTGAAGCACCAAGACTCCCTTTAGAGCATCACCTGTACGTAGTCACT 1848
 Qy 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
 Db 1849 TACGTAAATCTCGAGAGGTGACAAGGCTGACTGACCGGTGGCTACTTCACATTTCTTCTGC 1908
 Qy 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
 Db 1909 ATCAGTCAGCACTGTGATCTTTTATAGTATAGTATAGTATAGTATAGTATAGTATAGT 1968
 Qy 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
 Db 1969 GTGTCCCTTTTACAAGCTATCAAGTCTCTGAAGATGACCACTTGAACCAACAAAGGAATTT 2028
 Qy 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluGluIlePhe 620
 Db 2029 TGGGCCCACTTTTGGATTTCAGCAGGTCTCTCTCTGCTATATCTACTTACTCTCCCAAAATTTTC 2088

QY	621	SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu	640
Db	2089	TCCTTTTGAAGTACTACTGGATTTACATTTGATGGGATGCTCTACAAGCCCTCATGATCTA	2148
QY	641	GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu	660
Db	2149	CAGCCTTGGAAAGAAATATCCTACTGTGCTTTTATATATGTTGGTCTCTCAGGTGCAGTTG	2208
QY	661	ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly	680
Db	2209	GTCAATAATCCGGTTTAAAGGAGTCAAGTATTTCCCTTTGAATACCTCAGCCTCTCTAGGT	2268
QY	681	TyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly	700
Db	2269	TATGTGTTGTAGTGATACACAACAGGGGATCCTGTCCACCGAGGGCTTTAAATTTGAAGGC	2328
QY	701	AlaPheIleTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr	720
Db	2329	GCCTTTAAATATATAAATGGTCAATAGAAATTTGACGATCAGGTGGAGGACTCCAATAT	2388
QY	721	LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyr	740
Db	2389	CTAGCTTCTCGATATATTTTCATTTAGCTTTAGATCGTGTGGGCATCCACGGCTGGTCCCTAT	2448
QY	741	GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle	760
Db	2449	GGAGGATACCTCTCCCTGATGGCATTTATGCAAGGTCAATCTTGGGTTCCTATT	2508
QY	761	AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet	780
Db	2509	GCTTGGGGCCCCAGTCACTCTGTGGATCTTCTATATACAGGATACACGGACGTTATATG	2568
QY	781	GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu	800
Db	2569	GGTCACCCCTCACCAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCCATCGAACAGAA	2628
QY	801	LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal	820
Db	2629	AAGTTCCCTCTGAAACCAATCGTTTACTGCTTTTACATGGTTCCTCGATGAGAAATGTC	2688
QY	821	HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp	840
Db	2689	CATTTTGCACATACCAGTATATTACTGAGTGTATTTTATAGTCAGGGCTGGAAAGCCATATGAT	2748
QY	841	LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr	860
Db	2749	TTACAGATCTATCTCTCAGGAGAGACACAGCATAAAGATTTCTTGAATCGGAGAACATTAT	2808
QY	861	GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLys	880
Db	2809	GAACTGGATCTTTTGGCATCTACCTTCAAGAAACCTTGGATCAGTATTGCTCTCTATAAA	2868
QY	881	ValIle 882	
Db	2869	GTGATA 2874	
RESULT 9			
US-09-976-674-12			
; Sequence 12, Application US/09976674			
; Patent No. US20020115843A1			
; GENERAL INFORMATION:			
; APPLICANT: Qi, Steve			
; APPLICANT: Akinsanya, Karen			
; APPLICANT: Riviere, Pierre			
; APPLICANT: Junien, Jean-Louis			
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV			
; FILE REFERENCE: 70669			
; CURRENT APPLICATION NUMBER: US/09/976,674			
; CURRENT FILING DATE: 2001-10-12			
; PRIOR APPLICATION NUMBER: US 60/240,117			
; PRIOR FILING DATE: 2000-10-12			
; NUMBER OF SEQ ID NOS: 61			
; SOFTWARE: PatentIn version 3.1			

1054 AAAATCTTGAATTTCTATATGAAGAAATGATGAATCTGAGGTGGAATATTTTCATGTT 1113
Db
301 ThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThr 320
QY
1114 ACATCCCTCATGTTGGAAACACAGAGGGCAGATTCATCCGTTATCCCTAAACAGGTACA 1173
Db
321 AlaAsnProLysValThrPheLysMetSerGluLeuMetIleAspAlaGluGlyArgile 340
QY
1174 GCAATCTCTAAAGTCATTTTAAGATGTGAGAAATANTGATGTGCTGAAGGAAGATC 1233
Db
341 IleAspValIleAspLysGluLeuIleGlnProPheGluIlePheGluGlyValGlu 360
QY
1234 ATAGATGTCTAGATAGGAACATAATCAACCTTTTGAGATCTTATTTGAAGAGCTTGA 1293
Db
361 TyrIleAlaArgAlaGlyThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
QY
1294 TATATTGCCAGAGCTGGATGCACTCCCTGAGGGAAATATGCTTGGTGCATCTCTAGAT 1353
Db
381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
QY
1354 CGCTCCAGACTCCCTTACAGATAGTGTGATCTCACTGGAATATTTATCCAGTAGAA 1413
Db
401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
QY
1414 GATGATGTTATGGAAAGCAGAGACTCATTTGAGTCAGTGCCTGATTTCTGTGAGGCCACTA 1473
Db
421 IleIleTyrGluGluThrThrAspIleTrrIleAsnIleHisAspIlePheHisValPhe 440
QY
1474 ATTATCTATGAAGAAACACAGACATCTGGATAATATCATGACATCTTTTCATGTTTTT 1533
Db
441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
QY
1534 CCCCAAGTCAAGAGGAAATTTGAGTTTATTTTGGCTTGAATGCAAAACAGGTTTC 1593
Db
461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
QY
1594 CGTCATTTATCAAAATATCATCTATTTTAAAGGAAGCAATATAAAGATCCAGTGGT 1653
Db
481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluIleAlaIleThrSer 500
QY
1654 GGGCTGCCTGCTCCCAAGTGAATTTCAAGTGTCTTATCAAGAGAGAGATAGCAATATCCAGT 1713
Db
501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
QY
1714 CGTGAATGGGAAGTTCTTTGGCCGCGCATGGATCTAATATCAAGTTGATGAAGTCAGAAG 1773
Db
521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer 540
QY
1774 CTGGTATATTTTGAAGGCCACCAAGATCCCTTTTAGAGCATCACTGTAGTGTAGTCACT 1833
Db
541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
QY
1834 TACGTAAATCTGGAGAGGTGCAAGGCTGACTGACCGTGGCTACTCACATTTCTTGCTGC 1893
Db
561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
QY
1894 ATCAGTCAGCATGCTGACTCTTTTATAGTAAGTATAGTAAACCAGAGAATCCACACTGT 1953
Db
581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
QY
1954 GTGTCCCTTTTACAGCTATCAAGTCCCTGAAGATGACCCCACTTGCAGAAACAAAGGAATTT 2013
Db
601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePhe 620
QY
2014 TGGGCCACCAATTTTGGATTCAGCAGGTCTCTTCTGACTATATCTCTCCAGAAATTTTC 2073
Db
621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
QY
2074 TCTTTTGAAGTACTACTGATTTTACATTTGATGGATGCTCTCAAGAGCCTCATGATCTA 2133
Db
641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGly--ProGlnValGlnL 660
QY

2134 CAGCCTGGAAGAAATATCTACTGTCTCTCATATATGTTGGTCTCTCCTCAGGTGCAGT 2193
Db
660 euValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuG 680
QY
2194 TGGTGAATATTCGGTTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTAGCCTCTCTAG 2253
Db
680 lYtyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluG 700
QY
2254 GTTATGTGTGTGTAGTAGACACAAAGGGGATCCCTGTACCAGAGGCTTAAATTTGAAG 2313
Db
700 lYValPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnT 720
QY
2314 GGCCTTTTAATATATAAATGGTCAATAGAAATTCAGATCAGGTGGAAGGACTCCAAT 2373
Db
720 yzLeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerT 740
QY
2374 ATCTAGCTTCTCGATATGATTTTCACTTAGATCTGTGGCATCCACGGCTGTCTCT 2433
Db
740 yzGlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaI 760
QY
2434 ATGGAGGATACCTCTCCCTGATGGCATTAATGAGAGGTGAGATATCTTCAGGGTTGCTA 2493
Db
760 leAlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrM 780
QY
2494 TTGCTGGGGCCCGAGTCACTCTGTGGATCTTCTATGATACAGGTACACGGAACGTTATA 2553
Db
780 etGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaG 800
QY
2554 TGGGTCACTCCCTGACAGATGAACAGGGCTATTACTTAGCATCTGTGGCCATCAAGCAG 2613
Db
800 luLysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnV 820
QY
2614 AAAAGTTCCCTCTGACCAATCGTTTACTGCTCTTACATGGTTCCTCGATGAGAATG 2673
Db
820 alHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrA 840
QY
2674 TCCATTTTGACATACCATGATATATTTAGTGTCTTTTGTAGGGCTGGAAAGCCATATG 2733
Db
840 splLeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHis 860
QY
2734 ATTTACAGATCTATCTCTCAGAGAGACACACATAAGATTTCTTGAATCGGAGAACATT 2793
Db
860 yzGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuL 880
QY
2794 ATGAACCTGCATCTTTTGCACTACCTTCAAGAAACCTTGGATCACGTAITGCTGCTCTAA 2853
Db
880 yzValIle 882
QY
2854 AAGTGATA 2861
Db

RESULT 10
US-10-982-512-12
; Sequence 12, Application US/10982512
; Publication No. US20050059081A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/10/982,512
; PRIOR FILING DATE: 2004-11-05
; PRIOR APPLICATION NUMBER: US/09/976,674
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ'ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 4829
; TYPE: DNA
; ORGANISM: Homo sapiens

US-10-982-512-12

Alignment Scores:

Alid. No.: 0 Length: 4829
Score: 4680.00 Matches: 882
Percent Similarity: 99.8% Conservative: 0
Best Local Similarity: 99.8% Mismatches: 0
Query Match: 99.8% Indels: 2
DB: 9 Gaps: 0

US-10-825-632-1 (1-882) x US-10-982-512-12 (1-4829)

Qy 1 MetAlaAlaMetGluThrGluGlnLeuGluValGluIlePheGluThrAlaAspCys 20
Db 214 ATGGCAGCAGCAATGGAAACAGAACAGCTGGGTGGATATTTGAGATATTTGAAACTCGGACTGT 273
Qy 21 GluGluAenIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
Db 274 GAGGAGAAATATTGAATCACAGGATCGGCCCTAAATTTGGAGCCCTTTTATGTTGAGCGGTAT 333
Qy 41 SerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
Db 334 TCCTGGAGTCAGCTTAAAAAGCTGCTTGGCCGATACAGAAAAATATCATGGCTACATGATG 393
Qy 61 AlalysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
Db 394 GCTAAGGACCCACATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTTCA 453
Qy 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
Db 454 GACAGAACTATTACCTTGCCATGCTCGTGTGAGAACAGAGAAATATACACTGTTTATTTCT 513
Qy 101 GluIleProLysThrIleAenArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
Db 514 GAAATTCCTCAAACTATCATAGACGACGCTTAAATGCTCTCTTGGAGCCCTTTTGG 573
Qy 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140
Db 574 GATCTTTTCAGGCACACATGGACTATGGAAATGATTTCTTCAGATATATCAACAGGAAGTGA 633
Qy 141 GluArgLysArgIleGlyThrValGlyLeAlaSerTyrAspTyrHisGlnGlySerGly 160
Db 634 GAAAGAAACCATTTGGAACAGTCGGAAATTTGCTTTCAGATATATCAACAGGAAGTGA 693
Qy 161 ThrPheLeuPheGlnAlaGlySerGlyLeTyrHisValLysAspGlyGlyProGlnGly 180
Db 694 ACATTTCTGTTTCAAGCCGCTAGTGGAAATTTATCACGTAAAAAGATGGAGGCCCAAGGA 753
Qy 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
Db 754 TTTACGCCAACCAACCTTTAAGGCCCAATCTAGTGGAACTAGTTGTCCCAACATACGGATG 813
Qy 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
Db 814 GATCCMAAATTTATGCCCTGCTGATCCAGACTGGATTTGCTTTATACATAGCAACGATATT 873
Qy 221 TrpIleSerAenIleValThrArgGluGluArgLeuThrTyrValHisGlnLeu 240
Db 874 TGGATATCTAACATCTGTAACAGAGAAAGAAAGAGACTCACTTATGTGCAACATGAGCTA 933
Qy 241 AlaAenMetGluGluAenAlaArgSerAlaGlyValAlaThrPheValLeuGluGluGlu 260
Db 934 GCCAACATGGAAAGATGCGAGATCAGCTGGAGTGGCTTCTTTGTCTTCCAGAGAA 993
Qy 261 PheAspArgTyrSerGlyTyrTrpCysProLysAlaGluThrThrProSerGlyGly 280
Db 994 TTTGATAGATATTCCTGGCTATTTGGTGTGTCCAAAAGCTGAAACACTCCCAAGTGGTGT 1053
Qy 281 LysIleLeuArgIleLeuTyrGluGluAenAspGluSerGluValGluIleIleHisVal 300
Db 1054 AAAATTTCTTGAATTTCTATATGAAGAAAAATGATGAATCTCGAGGTGGAAATTTATTCATGTT 1113
Qy 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320

Db 1114 ACATCCCTATGTTGGAAACAAGGAGGCGCAGATTTCATTCGTTTATCCTAAAAACAGGTACA 1173
Qy 321 AlaAenProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
Db 1174 GCBAATCCTAAAGTCACATTTTAAGATGTAGAAATATATGATTCATGCTGAAGGAAGATC 1233
Qy 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
Db 1234 ATAGATGTCATAGATAGAGAACTAATTCACCTTTTGAGATTCTATTTTAGAGGTGAA 1293
Qy 361 TyrIleAlaArgAlaGlyTyrThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
Db 1294 TATATGGCAGAGCTGGATGGACTCTCGAGGAAAAATATCTTGTCCTACCTACTAGAT 1353
Qy 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
Db 1354 CGCTCCACAGACTCGCTACAGATAGTGTGATCTCACCTGAATATTATTATCCAGTAGAA 1413
Qy 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
Db 1414 GATGATGTTATGGAAAGCGCAGAGACTCAATTGAGTCACTGCTGATTCGTGACGCCACTA 1473
Qy 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
Db 1474 ATTATCTATGAAGAACCAACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTT 1533
Qy 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
Db 1534 CCCCAAGTCACGAAGAGAAATTTAGTTTATTTTGGCTCTCAATGCAAAAACAGGTTTC 1593
Qy 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
Db 1594 CGTCAATTTATCAAAAATTTACATCTATTTTAAAGGAAGCAAAATATAACGATCCAGTGT 1653
Qy 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluLeuIleThrSer 500
Db 1654 GGGCTGCTCTCCAGTGATTTCAAGTGCTCTATCAAGAGAGATAGCAATTACAGT 1713
Qy 501 GlyGluTrpGluValLeuGlyArgHisGlySerAenIleGlnValAspGluValArgArg 520
Db 1714 GGTGAATGGAGATCTTTGGCCGCGATGGATCTAATATCCAACTGATGAAGTCAGAGG 1773
Qy 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisIleHisLeuTyrValValSer 540
Db 1774 CTGGTATATTTTGAAGGCGCCAAAGACTCCCTTTAGAGCATCACCTGACGTAGTCACT 1833
Qy 541 TyrValAenProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
Db 1834 TACGTAAATCTCTGGAGAGGTGACAAAGCTGACTGACCGCTGCTACTACATTTCTTGCTGC 1893
Qy 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
Db 1894 ATCAGTCAGCACTGTGACTCTTTTATTAAGTATAGTAACTAACCCAGAAAGATCCCACTGT 1953
Qy 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
Db 1954 GTGTCTCTTTACAGCTATCAAGTCTCTGAAGATGACCCCACTTGCAAAAACAAAGGAATTT 2013
Qy 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProPheGluIlePhe 620
Db 2014 TGGGCCACCAATTTTGGATTCAGCAGGTCTCTCTCTGACTATACTCTCTCCAGAAATTTTC 2073
Qy 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
Db 2074 TCTTTTGAAGTACTACTGGATTTTACATTTATATGGAGTCTCTACAGGCTCATGATCTA 2133
Qy 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGly--ProGlnValGlnL 660
Db 2134 CAGCTCGAAAGAAATATCTTACTGTCTGTTCATATATGGTGGTCTCTCTCAGGTGCACT 2193
Qy 660 euValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAenThrLeuAlaSerLeuG 680

Qy	341	IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu	360
Db	1234	ATAGATGTCATAGATAAGGAACATAATCAACCTTTTGGAGATTCTATTGGAAGGAGTTGAA	1293
Qy	361	TyrIleAlaArgAlaGlyTyrThrProGluGlyLysTyrAlaTyrPheSerIleLeuLeuAsp	380
Db	1294	TATATTGCCAGAGCTGGATGGACTCTCTGAGGGGAAATATGCTTGGTCCATCCTACTAGAT	1353
Qy	381	ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu	400
Db	1354	CGCTCCAGACTCCCTCAGATAGTGTGATCTCACCTGATATTATTATCCAGTAGAA	1413
Qy	401	AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu	420
Db	1414	GATGATGTTATGGAAAGCGAGACTCATTTGAGTCACTGCTGATCTCTGACGCCACTA	1473
Qy	421	IleIleTyrGluGluThrThrAspIleTyrPheIleAsnIleHisAspIlePheHisValPhe	440
Db	1474	ATTATCTATGAAGAAACAAACAGACATCTGGATAAATATCCATGACATCTTTCATGTTTT	1533
Qy	441	ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe	460
Db	1534	CCCCAAAGTCCAGAGAGGAAATTCAGTTTATTTTGGCTCTGAAATGCAAAACAGGTTTC	1593
Qy	461	ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly	480
Db	1594	CGTCATTTATACAAAATATACATCTATTTTAAAGGAAAGCAAAATATAAACGATCCAGTGGT	1653
Qy	481	GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer	500
Db	1654	GGGCTCGCTGCTCCCAAGTGATTTTCAAGTGTCTTATCAAGAGAGAGATAGCAATATACCACT	1713
Qy	501	GlyGluTyrGluValLeuGlyValArgHisGlySerAsnIleGlnValAspGluValArgArg	520
Db	1714	GGTGAATGGGAAGTCTTTGGCCGCGATGGATCTAATATATCAAGTTGATGAAGTCAGAAAGG	1773
Qy	521	LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisIleLeuTyrValValSer	540
Db	1774	CTGGTATATTTTGAAGGACCAAAAGACTCTCCCTTTAGAGCATCACTCTGACGTAGTCAGT	1833
Qy	541	TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys	560
Db	1834	TACGTAAATCCCTGGAGAGGTGACCAAGGCTGACTACCGTGGCTACTACATTTCTTGCTGC	1893
Qy	561	IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys	580
Db	1894	ATCAGTCAAGCATGTGACTTCTTTATAGTAAGTATAGTAACCAAGAGAAATCCACTGT	1953
Qy	581	ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe	600
Db	1954	GTGTCCCTTTACAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAAAACAAAGGAATTT	2013
Qy	601	TyrAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe	620
Db	2014	TGGGCCACCAATTTGGATTTCAGT-----	2036
Qy	621	SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu	640
Db	2036	-----	2036
Qy	641	GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu	660
Db	2037	-----CCTCAGGTGCAAGTTG	2051
Qy	661	ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly	680
Db	2052	GTGAATAATCGGTTTAAAGAGTCAAGTATTTCCCTTGATACCTTAGCTCTCTAGGT	2111
Qy	681	TyrValValValIleAspAsnArgLysSerCysHisArgGlyLeuLysPheGluGly	700
Db	2112	TATGTGGTTGTAGTAGACAAACAGGGGATCCTGTACCCAGGGCTTAAATTTGAAGGC	2171
Qy	701	AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr	720
Db	2172	GCCTTTAAATATAAATGGTCAATAGAAATGACGATCAGGTGGAAGGACTCCAATAT	2231
Qy	721	LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyLleHisGlyTyrPheSerTyr	740
Db	2232	CTAGCTTCTCGATATGATTTTCATTGACTTAGATCTGTGTGGGANTCCACGGCTGGTCTAT	2291
Qy	741	GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle	760
Db	2292	GGAGGATACCTCTCCCTGATGGCATTAATGACAGAGTCAGATATCTTCAGGGTTCTATT	2351
Qy	761	AlaGlyAlaProValThrLeuTyrPheTyrAspThrGlyTyrThrGluArgTyrMet	780
Db	2352	GCTGGGGCCCGACGCTACTCTGTGGATCTTCTATGATACAGGATACACGGAACGTTATATG	2411
Qy	781	GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu	800
Db	2412	GGTCACCTGACCAGAAATGAACAGGGCTATTACTTACATGCTCTTACATGCTGAGATGTC	2471
Qy	801	LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal	820
Db	2472	AACTTCCCTCTGAAACCAATCGTTTACTTCTTACATGCTCTTACATGCTGAGATGTC	2531
Qy	821	HisPheAlaHisThrSerIleLeuLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp	840
Db	2532	CATTTTGCACATACAGTATATTACTGAGTTTTTTAGTCAGGCTGGAAGGACATATGAT	2591
Qy	841	LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr	860
Db	2592	TTACAGATCTATCTCTCAGGAGACACAGCATTAAGAGTTCTTGAATCGGAGAACATTAT	2651
Qy	861	GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeuLys	880
Db	2652	GAACTGCACTCTTTTGCACTACCTTCAAGAAACCTTGGATCAGTATTGCTGCTCTAATA	2711
Qy	881	ValIle 882	
Db	2712	GTGATA 2717	
RESULT 12			
US-10-982-512-22			
; Sequence 22, Application US/10982512			
; Publication No. US20050059081A1			
; GENERAL INFORMATION:			
; APPLICANT: Q1, Steve			
; APPLICANT: Akinsanya, Karen			
; APPLICANT: Riviere, Pierre			
; APPLICANT: Junien, Jean-Louis			
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV			
; FILE REFERENCE: 70669			
; CURRENT APPLICATION NUMBER: US/10/982,512			
; CURRENT FILING DATE: 2004-11-05			
; PRIOR APPLICATION NUMBER: US/09/976,674			
; PRIOR FILING DATE: 2001-10-12			
; PRIOR APPLICATION NUMBER: US 60/240,117			
; PRIOR FILING DATE: 2000-10-12			
; NUMBER OF SEQ ID NOS: 61			
; SOFTWARE: PatentIn version 3.1			
; SEQ ID NO 22			
; LENGTH: 4685			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-10-982-512-22			
Alignment Scores:			
Pred. No.:	0	Length:	4685
Score:	4385.50	Matches:	834
Percent Similarity:	94.6%	Conservative:	0
Best Local Similarity:	94.6%	Mismatches:	1
Query Match:	93.3%	Indels:	48
DB:	9	Gaps:	1

US-10-825-632-1 (1-882) x US-10-982-512-22 (1-4685)

QY 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
DB 214 ATGGCAGCAGCAATGCAACAGAACAGCTGGGTGTGAGATATTTGAAACCTGGGACTGT 273
QY 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
DB 274 GAGGAGAAATATTGAATACAGATCGGCTTAAATTTGGAGCCCTTTTATGTTGAGCCGAT 333
QY 41 SerTrpSerGlnLeuLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
DB 334 TCTCGAGTCACTTAAAGCTGCTTGGCCGATACAGAAAATATCATGCTACATGATG 393
QY 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
DB 394 GCTAAGGCACCATGATTTTCAATGTTGTGAAGAGGAATGATCAGATGACCTCATTTCA 453
QY 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
DB 454 GACAGAAATCTATTACCTTTGCCATGCTGTGGTGAGAACAGAGAAAATACACTGTTTATTCT 513
QY 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
DB 514 GAAATTCCTCAAACTATCAATAGACGACAGCTTTAATGCTCTCTTGGAAAGCCCTTTTG 573
QY 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg 140
DB 574 GATCTTTTCAGCAACACTGGACTATGGAATGATTTCTCGAGAAGAAGAACTATTAAAG 633
QY 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
DB 634 GAAAGAAAACGCAATGGAAACAGTCGGAATTCCTTTACCAATATCACCAAGGAAGTGA 693
QY 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
DB 694 ACATTTCTGTTTCAAGCCGGTAGTGGAATTTATCACGTAAAGATGAGGGCCACAAAGA 753
QY 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
DB 754 TTTAGCAACACCTTTTAAAGCCCAATCTAGTGGAACACTAGTTGTCCCAACATACCGATG 813
QY 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
DB 814 GATCCAAAATATTGCCCTGCTGATCCAGACTGGATGCTTTTATACATAGCAACGATATT 873
QY 221 TrpIleSerAsnIleValThrArgGluGluArgArgLeuThrTyrValHisAsnGluLeu 240
DB 874 TGGATATCTAAATCATCGTAACAGAGAAAGAGAGACTCATTATGTGCAATGAGCTA 933
QY 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu 260
DB 934 GCCAATGGAAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAAGAA 993
QY 261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280
DB 994 TTTGATAGATATCTGGCTATTGGTGTGTCGCAAGCTGAACCACTCCACAGTGGTGT 1053
QY 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300
DB 1054 AAAATCTTTAGAAATCTATATGAAGAAAATGATGAATCTGAGGTGGAAATATTATCATGTT 1113
QY 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320
DB 1114 ACATCCCTTATGTTGGAACCAAGGAGGCGAGATTCATTTCCTGTTATCTCAAAACAGGTACA 1173
QY 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
DB 1174 GCMAATCTTAAGTCACTTTTAAGATGTGCAAAATATGATGATGCTGAGGAAGAGATC 1233
QY 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
DB 1234 ATAGATGTCATAGTAAGGAACCTAATTCACCTTTTGAGATTTCTATTTGAAGAGTTGAA 1293

QY 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
DB 1294 TATATTGCCAGAGCTGGATGGACTCTGAGGAAAATATGTTGGTCCATCTACTAGAT 1353
QY 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
DB 1354 CGTCCACAGACTCGCTACAGATAGTGTGATCTCACCTGAATTTATTTATCCAGTAGAA 1413
QY 401 AspAspValMetGluArgGlnArgIleGluSerValProAspSerValThrProLeu 420
DB 1414 GATGATGTTATGGAAGGCAGAGACTCATTTGAGTCAGTGCCTGATTTCTGTGACGCCACTA 1473
QY 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
DB 1474 ATTATCTATGAAGAAACAAACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTT 1533
QY 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
DB 1534 CCCAAAGTCACGAAGAGGAAATGAGTTATTTTGGCTCTGAAATGCAAAACAGGTTTC 1593
QY 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
DB 1594 CGTCATTTATACAAAATTACATCTATTTTAAAGGAAACAAATATAAATACGATCCAGTGT 1653
QY 481 GlyLeuProAlaProSerAspPheLysCysProLysLysGluGluIleAlaIleThrSer 500
DB 1654 GGGCTGCTGCTCCAAAGTGATTTCAAGTGCTCTATCAAAGAGAGAGATAGCAATTAACAGT 1713
QY 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
DB 1714 GGTGAATGGGAAGTTCTTGGCCGCGATGGATCTAATATCCAAAGTTGATGAAGTCAGAAG 1773
QY 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer 540
DB 1774 CTGGTATATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACCTGTAGTAGTCACT 1833
QY 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
DB 1834 TAGCTAAATCTCGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACATTTCTTGTGTC 1893
QY 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
DB 1894 ATCAGTCAGCACTGTGACTTCTTTATAAGTAAATAGTATAGTACCAAGAGAAATCCACACTGT 1953
QY 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
DB 1954 GTGTCCCTTTACAGCTATCAAGTCTGAGATGACCCCACTTGCNAACAAACAGAAATTT 2013
QY 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 620
DB 2014 TGGGCCACCACTTTGGATTCACT----- 2036
QY 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
DB 2036 ----- 2036
QY 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 660
DB 2037 -----CCTCAGGTGCGATTG 2051
QY 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuLeuAsnThrLeuAlaSerLeuGly 680
DB 2052 GTGAATATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATATCCCTAGCTCTCTAGGT 2111
QY 681 TyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 700
DB 2112 TATGTGGTGTAGTAGACAAACAGGGGATCCTGTCAACGAGGCTTAAATTTGAAGGC 2171
QY 701 AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGluTyr 720
DB 2172 GCCTTTAAATATATAAATGGGTCAATAGAAATTCAGATCAGGTGGAGGAGTCAATAT 2231

QY 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyr 740
DB 2232 CTAGCTTCTCGATATGATTTTCATTTGACTTAGATCGTGTGGGCATCCAGCGCTGGTCCCTAT 2291
QY 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
DB 2292 GGAGGATACCTCTCCCTGATGGCATTAAATGACAGGTCCAGATATCTTCAGGGTTGCTATT 2351
QY 761 AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780
DB 2352 GCTGGGGCCCGCAGTCACTCTGTGGATCTTCTATGATACAGGATACACCGAAGCTTATATG 2411
QY 781 GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
DB 2412 GGTCAACCTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAA 2471
QY 801 LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
DB 2472 AAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGGTTTCTCGATGAGAATGTC 2531
QY 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
DB 2532 CATTTTGACATATCTCTGACAGTATATTACTGAGTTTCTTGTAGTGGGCTGGAAGCCATATGAT 2591
QY 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
DB 2592 TTACAGATCTATCTCTCAGGAGAGACACAGCTATGAGTTTCTGATTCGGAGAACATTAT 2651
QY 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLys 880
DB 2652 GAACTGCAATCTTTTGCACTACCTTCAAGAAACCTTTGGATACGATTTCCTGCTCTAAAA 2711
QY 881 ValIle 882
DB 2712 GTGATA 2717

RESULT 13

US-09-976-674-20
; Sequence 20, Application US/09976674
; Patent No. US20020115843A1
; GENERAL INFORMATION:
; APPLICANT: QI, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976.674
; CURRENT FILING DATE: 2001-10-12
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 20
; LENGTH: 4676
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-976-674-20

Alignment Scores:
Pred. No.: 0 Length: 4676
Score: 4385.00 Matches: 831
Percent Similarity: 94.2% Conservative: 0
Best Local Similarity: 94.2% Mismatches: 1
Query Match: 93.3% Indels: 51
DB: 3 Gaps: 1

US-10-825-632-1 (1-882) x US-09-976-674-20 (1-4676)

QY 1 MetAlaAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
DB 214 ATGGCAGCAGCAATGGAAACAGAACACAGCTGGGTGTGGAGATATTTGAACTCGCGGATCTGT 273

QY 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
DB 274 GAGGAGATATTTGAATCACAGGATCGGCCCTAAATTCGAGCCCTTTTATGTTGAGCGGTAT 333
QY 41 SerTrpSerGlnLeuLysLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
DB 334 TCTCGAGTCACTTAAAGAGCTGCTGCGGATACCGAATAATATCATGGCTACATGATG 393
QY 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
DB 394 GCTAAGGACCAACATGATTTTCATGTTTGTGAAAGGAATGATCCAGATGGACCTCATTTCA 453
QY 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
DB 454 GACAGAATCTATTACCTTGCCTATGCTGGTGAGAACAGAGAAATACATGTTTTATTCT 513
QY 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
DB 514 GAAATTCGCAAACTATCATATAGACGACGATCTTAATGCTCTCTTGGAGGCTCTTTTG 573
QY 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg 140
DB 574 GATCTTTTTCAGGCAACACATATGGAATGTATTCGAGAAGAAGAACTATTAAAGA 633
QY 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
DB 634 GAAAGAAAACGCAATTTGGAACAGTCCGAAATGCTTCTTACGATATATCACCAGGAAGTGA 693
QY 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
DB 694 ACATTTCTGTTTCAAGCCGCTAGTGGAAATTTATCCAGTAAAGATGAGGGCCACAAGGA 753
QY 181 PheThrGlnGlnProLeuAspProAsnLeuValGluThrSerCysProAsnIleAspMet 200
DB 754 TTTACGCAACACCTTTAAGGCCAATCTAGTGGAACTAGTTGTCCCAACATACCGATG 813
QY 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
DB 814 GATCCAAAATATATGCCCTGCTGATCCAGACTGGATGCTTTTATATCATACCAACGATATT 873
QY 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu 240
DB 874 TGGATATCTTAACATCGTAACCCAGAGAGAAAGGAGACTCATCTATGTGCAATAGACTA 933
QY 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu 260
DB 934 GCCAATCGAAAGAGATGCCAGATCAGCTGAGTGCCTACCTTTGTTCTCCCAAGAGAA 993
QY 261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280
DB 994 TTTGATAGATATTTCTGGCTATTGGTGTGTCGCTGCTGCTGCTGCTGCTGCTGCTGCT 1053
QY 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300
DB 1054 AAAATTTCTAGAAATCTATATGAGAAATGATGAAATCTGAGGTGGAAATTTATTCATGTT 1113
QY 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320
DB 1114 ACATCCCTATGTTGGAAACAGGAGGCGAGATTCATTCCTGTTATCTCTAAAACAGGTACA 1173
QY 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
DB 1174 GCAATCCATAAGTCACTTTAAGATGTCAGAAATATGATGATGCTGCTGAGGAGGATC 1233
QY 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
DB 1234 ATAGATGTCATAGATAAGGAACATAATTCACCTTTTGAGATCTTATTTGAGAGGTGAA 1293
QY 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
DB 1294 TATATTTGCCAGAGCTGGATGGACTCTCTGAGGAAATATGCTGTGCTCTCTACTAGAT 1353
QY 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400

```
Db 1354 CGCTCCAGACTCGCTACAGATAGTGTGATCTCCCTGAATATTTATCCAGTAGAA 1413
Qy 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
Db 1414 GATGATGTTATGGAAGGACAGAGACTCATTTAGTCAGTCGCTGATTCGTGAGCCACTA 1473
Qy 421 IleIleTyrGluGluThrThrAspIleTyrIleAsnIleHisAspIlePheHisValPhe 440
Db 1474 ATTATCTATGAAGAAACAACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTTT 1533
Qy 441 ProGlnSerHisGluGluGluIleGluPheIlePheIleAspSerGluCysIleThrGlyPhe 460
Db 1534 CCCCAAGTCACGAAGAGAAATAGAGTTATTTTGGCTCTGAATGCAAAAACAGGTTTC 1593
Qy 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
Db 1594 CGTCATTTATACAAATTTACATCTATTTTAAAGGAAGCAATATATAAGATCCAGTGGT 1653
Qy 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluIleAlaIleThrSer 500
Db 1654 GGCTGCTGCTCCCAAGTGATTTCAAGTGCTCTATCAAGAGAGATAGCAATTACCACT 1713
Qy 501 GlyGluThrGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
Db 1714 GGTGAATGGGAAGTCTTTGGCGGCATGGATCTAATATCCCAAGTTGATGAAGTCGAAGG 1773
Qy 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisIleLeuTyrValValSer 540
Db 1774 CTGGTATATTTGAAGGCACCAAGACTCCCTTTAGAGCATCACTGTACGTAGTCAGT 1833
Qy 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
Db 1834 TACGTAAATCTGGAGAGGTGACCAAGCTGACTGACCGTGGCTACTCACATTTCTTGCTG 1893
Qy 561 IleSerGlnHisCysAspPheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
Db 1894 ATCAGTCAGCAGCTGTGACTCTTTTAAAGTAAGTATAGTAAACCAAGAAATCCACACTGT 1953
Qy 581 ValSerLeuTyrLysLeuSerSerProLysAspProThrCysLysThrLysGluPhe 600
Db 1954 GTGTCCTTTTACAAGCTATCAAGTCTCGAAGATGACCCACTGTCGAAAAACAAAGGAATTT 2013
Qy 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 620
Db 2014 TGGCCACCATTTTGATTCAGCAGGCTCTCTCTGACTATCTCTCTCAGAAATTTTC 2073
Qy 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
Db 2074 TCTTTTGAAGTACTACTGGATTTACATTTGATGGATGCTCTACAAGCCTCATGATCTA 2133
Qy 641 GlnProGlyLysGlyTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 660
Db 2134 CAGCCTGGAAGAAATATCTCTACTGTGCTGTTCATATATGTTGGTGGTCG----- 2180
Qy 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680
Db 2180 ----- 2180
Qy 681 TyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 700
Db 2180 ----- 2180
Qy 701 AlaPheLysTyrLysMetGlyGlnIleGluLeuAspGlnValGluGlyLeuGlnTyr 720
Db 2181 -----GTCATAATAGAAATTCAGCATCAGCTGGAAGGACTCCCAATAT 2222
Qy 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyr 740
Db 2223 CTAGCTTCTCGATATGATTTTCATTGACTTAGATCGTGTGGGCATCCACGGCTGGTCTAT 2282
Qy 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
```

```
Db 2283 GGAGGATACCTCTCCCTGATGGCATTAATGCAGAGGTCAGATATCTTCAGGGTTGCTATT 2342
Qy 761 AlaGlyAlaProValThrLeuTyrIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780
Db 2343 GCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACTTATATG 2402
Qy 781 GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
Db 2403 GGTCACTCCCTGACCAAGATGAACAGGGCTATTACTAGGATCTGTGGCCATGCAAGCAAA 2462
Qy 801 LysPheProSerGluProAsnArgLeuLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
Db 2463 AGTTTCCCTCTGACACCAATCGTTTACTGCTCTTACATGTTTCTCGATGAGAATGTC 2522
Qy 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
Db 2523 CATTTTGCACATACACAGTATATTAAGTATTTTGTAGGGCTGGAAGCCATATGAT 2582
Qy 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
Db 2583 TTACAGATCTATCTCCAGAGAGACACAGCATAAGAGTTCTGTAATCGGAGAACATTAT 2642
Qy 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLys 880
Db 2643 GAATGCATCTTTTGCACACTACCTTCAGAAAACCTTGGATCAGCTATTTGCTGCTTAAA 2702
Qy 881 ValIle 882
Db 2703 GTGATA 2708
RESULT 14
US-10-982-512-20
; Sequence 20, Application US/10982512
; Publication No. US20050059081A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPP1V
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/10/982,512
; CURRENT FILING DATE: 2004-11-05
; PRIOR APPLICATION NUMBER: US/09/976,674
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
; LENGTH: 4676
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-982-512-20
Alignment Scores:
Pred. No.: 0 Length: 4676
Score: 4385.00 Matches: 831
Percent Similarity: 94.2% Conservative: 0
Best Local Similarity: 94.2% Mismatches: 1
Query Match: 93.3% Indels: 51
DB: 9 Gaps: 1
US-10-825-632-1 (1-882) x US-10-982-512-20 (1-4676)
Qy 1 MetalAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
Db 214 ATGCAGCAGCAATGGAACAGACAGCTGGGTGTGAGATATTTGAAACTGCGGACTGT 273
Qy 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
Db 274 GAGGAGAATATGAATCACAGGATCGGCTAAATGGAGCCCTTTTATGTTGAGCGGTAT 333
```

QY 41 SerTrpSerGlnLeuLysLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
DB 334 TCCGAGAGTCAGCTTAAAGAGCTGCTTGGCGATACGAGAAATATCATGGCTACATGATG 393
QY 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
DB 394 GCTAAGGCACACATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTC 453
QY 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
DB 454 GACAGAAATCTATTACCTTGGCATGCTCTGGTGGAGAACAGAGAAAAATACATCTGTTTATCT 513
QY 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
DB 514 GAAATTCCTCAAAACATATCAATAGACGACAGTCTTAATGCTCTCTTGGAAAGCCTCTTTTG 573
QY 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg 140
DB 574 GATCTTTTTCAGGCAACACTGGACTATGGAAATGTAATCTCGAGAGAAAGAACTATTAA 633
QY 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
DB 634 GAAAGAAACCATGGACAGTCGNAATGCTTCTTACGATATACACCAAGGAAGTGA 693
QY 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
DB 694 ACATTTCTGTTTCAAGCGGTGAGTGGAAATTTATCAGTAAAGATGGAGGCCACAGGA 753
QY 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
DB 754 TTTAGCGCAACACCTTTAAGGCCCAATCTAGTGGAAACTAGTGTGCCCCAACATACGGATG 813
QY 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
DB 814 GATCCAAATTAATGCCCTGCTGATCCAGACTGGATTTGCTTTATACATAGCAACGATAT 873
QY 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisGlnLeu 240
DB 874 TGCATATCTAACATCGTAACACAGAGAGAAAGAGAGACTCACATTATGTGACCAATAGCTA 933
QY 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu 260
DB 934 GCCAACATGGAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGCTCTCCAAAGAA 993
QY 261 PheAspArgTyrSerGlyTyrTrpCysProLysAlaGluThrThrProSerGlyGly 280
DB 994 TTTGATAGATATCTGGCTATTGGTGGTGTCCAAAGCTGAACAACTCCACAGTGGTGGT 1053
QY 281 LysIleLeuArgIleLeuTyrGluLysAsnAspGluSerGluValGluIleHisVal 300
DB 1054 AAAATCTTTAGAAATTTCTATATGAAGAAATATGATGAATCTGAGGTGGAAATTTATCATGTT 1113
QY 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320
DB 1114 ACATCCCTATGTTGGAAACAGAGGCGCAGATTCATTCGTTATCTTAAACAGGTACA 1173
QY 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluIleArgIle 340
DB 1174 GCAATCTCCTAAAGTCACCTTTTAAAGATGTCAGAAATATATGATGATCTGAGAGAGGATC 1233
QY 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
DB 1234 ATAGATGTCATAGATAAGGAACTAAATTCACCTTTTGAAGATCTCATTTGAAGAGTTGAA 1293
QY 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
DB 1294 TATATTCGACAGCTGGATGACTCTCTGAGGAGAAATATGCTTTGGTCCATCTCTACTAGAT 1353
QY 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
DB 1354 CGCTCCACACTCGCCTACAGATAGTGTGATCTCACCTGAATTAATTTATCCAGGTAGAA 1413
QY 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420

DB 1414 GATCATGTTATGGAAGCGCAGAGACTCATTGAGTCAGTCGCTGATTTCTGTGACGCACTA 1473
QY 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
DB 1474 ATTATCTATGAAGAAACACACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTTT 1533
QY 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
DB 1534 CCCCAGAGTCAGAGAGAGGAAATTTGAGTTTATTTTGGCTCTGAATGCAAAACAGGTTTC 1593
QY 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerGly 480
DB 1594 CGTCATTTATACAAAATTTACATCTATTATTAAGAGAAAGCAAAATATTAACGATCCAGTGGT 1653
QY 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 500
DB 1654 GGGCTGCTGCTCCAAAGTGATTTCAAGTGTCTTATCAAGAGAGAGATAGCAATTTACCAGT 1713
QY 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
DB 1714 CGTGAATGGGAAGTTCTTGGCCGCATGGATCTAATATCCAAAGTTGATGAAGTCAAGAG 1773
QY 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer 540
DB 1774 CTGGTATATTTTGAAGGCACCAAGACTCTCCCTTTAGAGCATCACCTGTACGTAGTCACT 1833
QY 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
DB 1834 TACGTAATCTCGAGAGGTGCACAAAGCTGACTGACCGTGGCTACTCACATTTCTGCTGC 1893
QY 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
DB 1894 ATCAGTCAGCAGCTGTGACTTCTTTAAGTAAGTATAGTAACCAAGAAATCCACACTGT 1953
QY 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
DB 1954 GTGTCTCTTACAGACTATCAAGTCTCTGAGATGACCCAACTTGCAAAACAAAGGAATTT 2013
QY 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePhe 620
DB 2014 TGGGCCACATTTTGGATTCAGCAGGTCCTCTCTGACTATATCTCTCCGAAATTTTC 2073
QY 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
DB 2074 TCTTTTGAAGTACTACTGGATTTTACATTTGTATGGGATGCTCTACAGCCTCATGATCTA 2133
QY 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 660
DB 2134 CAGCCTGGAAGAAATATCTACTGCTGTGTTTCATATATGTTGCTCG----- 2180
QY 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680
DB 2180 ----- 2180
QY 681 TyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 700
DB 2180 ----- 2180
QY 701 AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr 720
DB 2181 -----GGTCAATAGAAATTCAGCATCAGGTGGAAGACTCCAAAT 2222
QY 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyr 740
DB 2223 CTAGCTTCTCGATATGATTTTCATGACTAGATGCTGTGGGCATCCACGGCTGCTCTAT 2282
QY 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
DB 2283 GGAGGATACCTCTCCCTGATGGCATTAATGACAGAGTCAGATACTTCAGGGTTGCTATT 2342
QY 761 AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780

```
Db 2343 GCTGGGGCCAGTCACCTCTGTGGATCTTCTATGATACAGATACACGGACGTTATATG 2402
Qy 781 GlyHisProAspGlnAsnGlnGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
Db 2403 GGTCAACCTGACCAAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAA 2462
Qy 801 LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAnVal 820
Db 2463 AAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGGTTCTCTGGATGAGAATGTC 2522
Qy 821 HisPheAlaHisThrSerLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
Db 2523 CATTTGGACATACAGATATATTACTGATTTTTTTAGTGGGCTGGAAAGCCATATGAT 2582
Qy 841 LeuGlnIleTyrProGlnGlnArgHisSerIleArgValProGluSerGlyGluHisTyr 860
Db 2583 TTACGATCTATCTCTCAGGAGACACAGCATAAGAGTTCTGAAATCGGAGAACATTAT 2642
Qy 861 GluLeuHisLeuLeuHisTyrLeuGlnGlnAsnLeuGlySerArgIleAlaAlaLeuLys 880
Db 2643 GAACCTGCATCTTTTGCACCTACCTTCAAGAAAAACCTTGGATCACGTTATGCTGCTTAAAA 2702
Qy 881 ValIle 882
Db 2703 GTGATA 2708

RESULT 15
US-10-956-157-2177
; Sequence 2177, Application US/10956157
; Publication No. US20050118625A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
; FILE OF INVENTION: HUMAN OSTEOARTHRITIS AND HUMAN PROTEASES
; FILE REFERENCE: 031896-043000 (AM 101081)
; CURRENT APPLICATION NUMBER: US/10/956,157
; NUMBER OF SEQ ID NOS: 319805
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2177
; LENGTH: 2830
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-956-157-2177

Alignment Scores:
Pred. No.: 0 Length: 2830
Score: 4118.00 Matches: 782
Percent Similarity: 88.7% Conservative: 0
Best Local Similarity: 88.7% Mismatches: 0
Query Match: 87.6% Indels: 100
DB: 9 Gaps: 1

US-10-825-632-1 (1-882) x US-10-956-157-2177 (1-2830)
Qy 1 MetAlaAlaMetGluThrGluLeuGlnGlyValGluIlePheGluThrAlaAspCys 20
Db 214 ATGGCAGCAGCAATGGAACAGACAGCCTGGGTGTGAGATATTTGAAATCTGCGACTGT 273
Qy 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
Db 274 GAGGAGAAATTTGAATCACAGGATCGGCTAAATTTGGAGCCTTTTATGTTGAGCGGTAT 333
Qy 41 SerTrpSerGlnLeuLysLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
Db 334 TCCTGGAGTCAGCTTAAAAAGCTGCTTCCGATACCAAGAAATATCATCATGCTCATGATG 393
Qy 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
Db 394 GCTAAGGCACCAATGATTTTCATGTTTGTGAAGAGGAATGATTCAGATGACCTCATCA 453
Qy 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
```

```
Db 454 GACAGAAATCTATTACCTTGCCATGTCGTGTGAGAACAGAGAAAATAACACTGTTTATCT 513
Qy 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
Db 514 GAAATTTCCCAAACTATCAATAGACAGCAGCTCTTAATGCTCTCTTGGAAGCCTCTTTTG 573
Qy 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg 140
Db 574 GATCTTTTTCAGGCAACACTGGACTATGGAATGATTTCTCAGAGAAGAACTATTAAAGA 633
Qy 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
Db 634 GAAAGAAAACGATATGGAAACAGTCGGAATTCCTTACGATTATCACCAGAGGAAGTGA 693
Qy 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
Db 694 ACATTTCTGTTTCAAGCGGTAGTGGAAATTTATCAGCTAAAGATGAGGCGCCACAAGA 753
Qy 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
Db 754 TTTTACGCAACACCTTTTAGGCCCAATCTAGTGGAAACTAGTTCTGCCAACATACGATG 813
Qy 201 AspProLysLeuCysProAlaAspProAlaAspTrpIleAlaPheIleHisSerAsnAspIle 220
Db 814 GATCCAAAATTTATGCCCTGCTGATCCAGACTGGATGCTTTTATACATAGCACGATATT 873
Qy 221 TrpIleSerAsnIleValThrArgGluGluArgArgLeuThrTyrValHisAsnGluLeu 240
Db 874 TGGATATCTAACATCGTAACACAGAGAGAAGAGGACTCACTTATGTGCACAATGAGCTA 933
Qy 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu 260
Db 934 GCCAACATGGAAGAAGATCCAGATCAGCTGGAGTCGCTACTTTTGTCTCCAAAGAAGA 993
Qy 261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280
Db 994 TTTGATAGATATTTCTGGCTATTGGGTGGTGTCCAAAGCTGAACAACTCCACGFGTGGT 1053
Qy 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300
Db 1054 ABAATTTCTAGAAATCTATATGAGAAATGATGAATCTGAGGTGGAAATTTATTCATGTT 1113
Qy 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320
Db 1114 ACATCCCTATGTTGGAAACAAGAGGCGCAGATTCAATTCCTGTTATCTCTAAACAGGTACA 1173
Qy 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
Db 1174 GCAAAATCCTTAAAGTCACTTTTAAAGATGTCAGAAATAATGATGTGCTGAAGGAAGATC 1233
Qy 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
Db 1234 ATGATGTCTATAGATAGGAAGAACTAATTCACCTTTTGAGATCTTATTTGAGGAGTTGAA 1293
Qy 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
Db 1294 TATATTGCCAGAGCTGGATGGACTCTCTGAGGGAATAATGCTGTGCTCATCTACTAGAT 1353
Qy 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
Db 1354 CGCTCCAGACTGCGCTACAGATAGTGTGATCTCACCTGAAATTTATTTATCCCAAGTGA 1413
Qy 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
Db 1414 GATGATGTTATGGAAGGAGGAGACTCATTTGATGCTGCTGCTGATCTGTGACGCCACTA 1473
Qy 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
Db 1474 ATTATCTATGAGAAAACAACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTT 1533
Qy 441 ProGlnSerHisGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
```

```

Db 1534 CCCAAGTACGAGAGGAATTTGAGTTATTTTGGCTCGAATGCAAAACAGGTTTC 1593
Qy 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
Db 1594 CGTCATTTATACAAAATTTACATCTATTTAAAGGAAGCAAAATAAAACGATCCAGTGGT 1653
Qy 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 500
Db 1654 GGGCTCGCTGCTCCAGTGATTTCAAGTGTCTCTATCAAAAGAGGAGATAGCAATACCAGT 1713
Qy 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArg 520
Db 1714 GGTGAATGGGAAGTTCTTTGGCGGCATGGATCTAATATCCAAAGTTGATGAAGCAGAGG 1773
Qy 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer 540
Db 1774 CTGGTATATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTAGTAGTCAGT 1833
Qy 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
Db 1834 TACGTAAATCTCGAGAGGTGACNAGGCTGACTGACCGTGGCTACTCACATTTCTTGCTGC 1893
Qy 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
Db 1894 ATCAGTCAGCACTGTGACTCTTTATAGTAGTATAGTATACCAAGAGNAATCCACACTGT 1953
Qy 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
Db 1954 GTGTCCTTTTAAAGCTATCAAGTCTGGAAGATGACCCAACTTGCAAAACAAAGGAATTT 2013
Qy 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 620
Db 2014 TGGGCGACCAATTTTGGATTTCAGCAGGTCCTCTTCTGACTATACTCTCCAGAAATTTTC 2073
Qy 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
Db 2074 TCTTTTGAAGTACTACTGGATTTCATTTGTATGGATGCTCTACAAAGCCTCATGATCTA 2133
Qy 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 660
Db 2134 CAGCCTGGAAGAAATATCTTACTGTGCTGTCTCATATATGGTGGTCTCTCAG----- 2184
Qy 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680
Db 2184 ----- 2184
Qy 681 TyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 700
Db 2184 ----- 2184
Qy 701 AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr 720
Db 2184 ----- 2184
Qy 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyr 740
Db 2184 ----- 2184
Qy 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
Db 2185 -----GTTGCTATT 2193
Qy 761 AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780
Db 2194 GCTGGGGCCCCAGTCCTGTGGATCTTCTATGATACAGGATACACGGAACGTTATATG 2253
Qy 781 GlyHisProAspGlnAsnGluGlnGlyTyrThrLeuGlySerValAlaMetGlnAlaGlu 800
Db 2254 GGTCAACCTTGACCAAGNAATGACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAA 2313
Qy 801 LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
Db 2314 AGTTTCCCTCTGACCAAAATCGTTTACTGTCTTACATGGTTTCTTGGATGAGATGTC 2373

```

```

Qy 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
Db 2374 CATTTGCACATACCACTATATTACTGAGTTTTTTAGTCAGGGCTGGAAAGCCATATGAT 2433
Qy 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
Db 2434 TTACAGATCTATCTCTCAGGAGACACACAGCATAAGAGTTCTCTGAATCCGGAGAACATTAT 2493
Qy 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLys 880
Db 2494 GAACCTGCATCTTTTGCACCTACCTTCAAGAAAACCTTGGATCTGCTCTCTAAAA 2553
Qy 881 ValIle 882
Db 2554 GTGATA 2559

```

Search completed: April 15, 2006, 01:36:23
Job time : 1586 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

- OM protein - nucleic search, using frame_plus_p2n model

Run on: April 15, 2006, 01:11:56 ; Search time 855 Seconds
(without alignments)
4155.026 Million cell updates/sec

Title: US-10-825-632-1

Perfect score: 4700

Sequence: 1 MAAAMTEQLGVIEFTADC.....HLHYLOENLGSRIAALKVI 882

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 9281099 seqs, 2013915447 residues

Total number of hits satisfying chosen parameters: 18562198

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=slp
-Q=/abs/ABSSWEB.spool/US10825632/runat_14042006_092613_5040/app_query.fasta_1
-DB=Published Applications NA New -QFMT=faastp -SUFFIX=p2n.rnpbn -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -HOST=abs05p
-USER=US10825632 @CGN 1.1 991 @runat_14042006_092613_5040 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA New:

1: /SID55/ptodata/2/pubpna/US08_NEW_PUB.seq.*
2: /SID55/ptodata/2/pubpna/US06_NEW_PUB.seq.*
3: /SID55/ptodata/2/pubpna/US07_NEW_PUB.seq.*
4: /SID55/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
5: /SID55/ptodata/2/pubpna/US09_NEW_PUB.seq.*
6: /SID55/ptodata/2/pubpna/US09_NEW_PUB.seq.*
7: /SID55/ptodata/2/pubpna/US10_NEW_PUB.seq.*
8: /SID55/ptodata/2/pubpna/US10_NEW_PUB.seq.*
9: /SID55/ptodata/2/pubpna/US10_NEW_PUB.seq.*
10: /SID55/ptodata/2/pubpna/US10_NEW_PUB.seq.*
11: /SID55/ptodata/2/pubpna/US11_NEW_PUB.seq.*
12: /SID55/ptodata/2/pubpna/US11_NEW_PUB.seq.*
13: /SID55/ptodata/2/pubpna/US11_NEW_PUB.seq.*
14: /SID55/ptodata/2/pubpna/US11_NEW_PUB.seq.*
15: /SID55/ptodata/2/pubpna/US60_NEW_PUB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	4700	100.0	2649	14	US-11-151-601-21
2	4700	100.0	3143	14	US-11-151-601-19
					Sequence 21, Appl
					Sequence 19, Appl

3	651	13.9	2238	11	US-11-079-463-1186	Sequence 1186, Ap
4	551	11.7	1884	11	US-11-079-463-2536	Sequence 2536, Ap
5	529	11.3	2217	11	US-11-208-288-3	Sequence 3, Appl
6	529	11.3	3332	11	US-11-208-288-1	Sequence 1, Appl
7	529	11.3	3407	9	US-10-501-035-34	Sequence 34, Appl
8	525	11.2	2301	8	US-10-522-789-1	Sequence 1, Appl
9	517.5	11.0	2283	11	US-11-208-288-5	Sequence 5, Appl
10	490	10.4	2778	9	US-10-932-182A-5649	Sequence 5649, Ap
11	490	10.4	2778	9	US-10-932-182A-5649	Sequence 5649, Ap
12	478.5	10.2	4852	14	US-11-136-527-2130	Sequence 2130, Ap
13	454	9.7	2814	14	US-11-245-147-168	Sequence 168, Ap
14	454	9.7	2814	14	US-11-186-284-54	Sequence 54, Appl
15	436.5	9.3	2457	9	US-10-932-182A-1107	Sequence 1107, Ap
16	436.5	9.3	2457	9	US-10-932-182A-1107	Sequence 1107, Ap
17	433	9.2	2819	14	US-11-136-527-2716	Sequence 2716, Ap
18	278.5	5.9	535	11	US-11-226-869-428	Sequence 428, Ap
19	252.5	5.4	1875	11	US-11-079-463-2282	Sequence 2282, Ap
20	215	4.6	1346	6	US-09-925-065A-669313	Sequence 669313, A
21	211.5	4.5	600	14	US-11-136-527-6812	Sequence 6812, Ap
22	206.5	4.4	2214	11	US-11-079-463-4059	Sequence 4059, Ap
23	200	4.3	609	6	US-09-925-065A-743558	Sequence 743558, A
24	200	4.3	609	6	US-09-925-065A-743558	Sequence 743558, A
25	194	4.1	378	9	US-10-932-182A-81332	Sequence 81332, A
26	194	4.1	378	9	US-10-932-182A-81332	Sequence 81332, A
27	181.5	3.9	354	9	US-10-932-182A-173537	Sequence 173537, A
28	181.5	3.9	354	9	US-10-932-182A-173537	Sequence 173537, A
29	156	3.3	2384	14	US-11-136-527-2023	Sequence 2023, Ap
30	130.5	2.8	2112	9	US-10-932-182A-81333	Sequence 81333, A
31	130.5	2.8	2112	9	US-10-932-182A-81333	Sequence 81333, A
32	127	2.7	612	6	US-09-925-065A-818446	Sequence 818446, A
33	126.5	2.7	4344	9	US-10-932-182A-3542	Sequence 3542, Ap
34	126.5	2.7	4344	9	US-10-932-182A-3542	Sequence 3542, Ap
35	125.5	2.7	5493	9	US-10-932-182A-81749	Sequence 81749, A
36	125.5	2.7	5493	9	US-10-932-182A-81749	Sequence 81749, A
37	125	2.7	4257	9	US-10-932-182A-4976	Sequence 4976, Ap
38	125	2.7	4257	9	US-10-932-182A-4976	Sequence 4976, Ap
39	124	2.6	9023	8	US-10-995-561-134	Sequence 134, App
40	123.5	2.6	1694969	7	US-10-506-454-1690	Sequence 1690, Ap
41	121.5	2.6	3464	8	US-10-995-561-465	Sequence 465, App
42	119.5	2.5	3387	14	US-11-091-883-183	Sequence 183, App
43	119.5	2.5	3439	14	US-11-000-688-851	Sequence 851, App
44	118.5	2.5	3468	14	US-11-098-686-9205	Sequence 9205, Ap
45	118.5	2.5	3729	8	US-10-453-372-1135	Sequence 1135, Ap

ALIGNMENTS

RESULT 1

US-11-151-601-21
; Sequence 21, Application US/11151601
; Publication No. US2006003413A1

GENERAL INFORMATION:

; APPLICANT: Millennium Pharmaceuticals, Inc.

; APPLICANT: Meyers, Rachel E.

; APPLICANT: Olandt, Peter J.

; APPLICANT: Kapeller-Libermann, Rosana

; APPLICANT: Curtis, Rory A. J.

; APPLICANT: Williamson, Mark

; APPLICANT: Welch, Nadine

; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASE, PHOSPHATASE,

; TITLE OF INVENTION: AND PROTEASE FAMILY MEMBERS AND USES THEREOF

; FILE REFERENCE: MPI00-054PIRCPIOMNIDVIM

; CURRENT APPLICATION NUMBER: US/11/151,601

; PRIOR FILING DATE: 2005-06-13

; PRIOR APPLICATION NUMBER: US 10/170,789

; PRIOR FILING DATE: 2002-06-13

; PRIOR APPLICATION NUMBER: US 09/797,039

; PRIOR FILING DATE: 2001-02-28

; PRIOR APPLICATION NUMBER: PCT/US01/06525

; PRIOR FILING DATE: 2001-02-28

; PRIOR APPLICATION NUMBER: US 60/186,061

; PRIOR FILING DATE: 2000-02-29

; PRIOR APPLICATION NUMBER: US 09/882,166

```

; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: PCT/US01/19269
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/212,078
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/934,406
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: PCT/US01/26052
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/226,740
; PRIOR FILING DATE: 2000-08-21
; Remaining prior application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 2649
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-11-151-601-21

Alignment Scores:
Pred. No.: 0 Length: 2649
Score: 4700.00 Matches: 882
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 14 Gaps: 0

US-10-825-632-1 (1-882) x US-11-151-601-21 (1-2649)
QY 1 MetAlaAaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
DB 1 ATGGCAGCAGCAATGGAAACAGAACAGCTGGGTGTTGAGATATTGGAATCTCGGACTGT 60
QY 21 GluGluAenIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
DB 61 GAGGAGATATTGGAATCATCAGATCGGCTAAATTTGGAGCTTTTATGTTGAGCGGTAT 120
QY 41 SerTrpSerGlnLeuLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
DB 121 TCTGGAGTCAGCTTAAAGCTGCTTGGCGATACCAAGAAATATCATGGCTACATGATG 180
QY 61 AlaLysAlaProHisAspMetPheValLysArgAsnAspProAspGlyProHisSer 80
DB 181 GCTAAGCCACCATGATTCATGTTTGTGAAGAGGAATGATCCAGATGGAACCTCATTTCA 240
QY 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAenThrLeuPheTyrSer 100
DB 241 GACGAAATCTATTATTCCTTGGCATGCTGTGAGAACAGAGAAATACACTGTTTATTCT 300
QY 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
DB 301 GAAATTTCCAAACATCATATAGACAGCAGCTTATGCTCTCTTGAAGCCTCTTTTG 360
QY 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140
DB 361 GATCTTTTTCAGCAACATCGACTATGGAATGATATCTCGAAGAAAGAACTATTAAAG 420
QY 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
DB 421 GAAAGAAACCGCATGGAAACATCGGAATGCTTCTACGATATATACCAAGAGGTGGA 480
QY 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
DB 481 ACATTTCTGTTTCAAGCCGGTAGTGGAAATTTATCATCGTAAAGAGATGAGGGCCACAGGA 540
QY 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
DB 541 TTTACGCAACAACTTTAAGGCCAATCTAGTGGAAACTAGTTGTCCTCCAAACATCGGATG 600
QY 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
DB 601 GATCCAAAATTAATGCTTGTGATCCAGACTGGATTGCTTTTATATACATAGCAACGATATT 660

```

```

QY 221 TrpIleSerAsnIleValThrArgGluGluArgArgLeuThrTyrValHisAsnGluLeu 240
DB 661 TGGATATCTAAACATCGTAAACAGAGAAAGAGAGACTTACTTATGTGCACAATGAGCTA 720
QY 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaAlaThrPheValLeuGlnGlu 260
DB 721 GCCAACATGGAAGAAGATGCCAGATCAGCTGGAGTCGTACCTTTGTTCTCCAAAGAA 780
QY 261 PheAspArgTyrSerGlyTyrTrpTyrCysProLysAlaGluThrThrProSerGly 280
DB 781 TTTGATAGATATTCTGGCTATTCTGGTGTCTCCAAAGCTGAAACCACTCCCAAGTGTGT 840
QY 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300
DB 841 AAAATTTCTTGAATTTCTATATGAAGAAATGATGAATCTGAGGTGGAATATTCTGTT 900
QY 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320
DB 901 ACATCCCTATGTTGGAACCAAGGAGCGCAGATTCATTCGGTTATCTCTAAACACAGTACA 960
QY 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
DB 961 GCAAATCTCTAAAGTCACTTTTAAAGATGTCAGAAATATGATGATGCTGAAGGAAGGATC 1020
QY 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
DB 1021 ATAGATGTCATAGATAAGGAACTAAATCAACCTTTTGGAGATCTCTATTTGAAGGAGTTGAA 1080
QY 361 TyrIleAlaArgAlaGlyTyrThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
DB 1081 TATATTGCCAGAGCTGGATGGACTCTCTGAGGAAATATGCTTGGTCCATCTCTACTAGAT 1140
QY 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
DB 1141 CGCTCCAGACTCGCTCGCATAGTAGTGTGATCTCACCTGAAATTTATTTATCCAGTAGAA 1200
QY 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
DB 1201 GATGATGTTATGGAAGGAGAGACTCATTTGAGTCAGTCGCTGATCTGTGAGCCACTA 1260
QY 421 IleIleTyrGluGluThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
DB 1261 ATTATCTATGAAGAAACCAACAGACATCTGGNTAATATCCATGACATCTTTTCATGTTT 1320
QY 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
DB 1321 CCCCAAGTCACGAAGAGAAATGAGTTTATTTTTCCTCTGATGCAAAAACAGGTTTC 1380
QY 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
DB 1381 CGTCATTTATACAAAATTTACATCTATTTTAAAGGAAAGCAATATAAACGATCCAGTGGT 1440
QY 481 GlyLeuProAlaProSerAspPheLysCysProLysGluGluIleAlaIleThrSer 500
DB 1441 GGGCTGCTGCTCCAAAGTATTTCAAGTGTCTCTATCAAGAGGAGATAGCAATTTACCAGT 1500
QY 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
DB 1501 GGTGAATGGAAAGTCTTGGCCGCGCATGGATCTAATATCCAAAGTTGATGAAGTCAGAAGG 1560
QY 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer 540
DB 1561 CTGGTATATTTTGAAGGCCACCAAGACTCCCCCTTTAGAGCATACCTGTGTAGTGTAGT 1620
QY 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
DB 1621 TACGTAAATCTCTGAGAGGCTGACAAAGGCTGACCTGACCGTGGCTACTCACATTTCTGTGC 1680
QY 561 IleSerGlnHisCysAspPheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
DB 1681 ATCAGTCAGCAGCTGTGACTTCTTTTATAGTAAGTATAGTAACCAAGAGAAATCCACACTGT 1740

```



```
Qy 581 ValSerLeuTyrIleuSerSerProGluAspPProThrCysLysThrLysGluPhe 600
Db 1741 GTGTCCTCTTTACAGCTATCAAGTCTCTCAAGATGACCACTTGCAAAACAAAGGAATTT 1800
Qy 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 620
Db 1801 TGGGCCACCAATTTGGATTGACAGGTCTCTTCTGACTATATCTCTCCAGAAATTTTC 1860
Qy 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
Db 1861 TCTTTTGAAGTACTACTGGATTTACATTTGATGGATGCTCTACAGCCCTCATGATCTA 1920
Qy 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 660
Db 1921 CAGCCTGGAAAGAAATATCTACTGCTGTGTTTCATATATGCTGCTCAGGTGCAAGTTG 1980
Qy 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680
Db 1981 GTGATATATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCCCTAGCTCTCTAGGT 2040
Qy 681 TyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 700
Db 2041 TATGTGTTGTAGTAGACACAACAGGGGATCTGTCCACCGAGGGCTTAAATTTGAAGGC 2100
Qy 701 AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr 720
Db 2101 GCCTTTAAATATAAATGGGTCAATAGAAATTCAGCATCAGGTGGGAAGGACTCCCAATAT 2160
Qy 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyr 740
Db 2161 CTAGCTTCTCGATATGATTTTCATCTGACTTAGATGCTGGGCATCCACGGCTGGTCTAT 2220
Qy 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
Db 2221 GGGAGGATACCTCTCCCTGATGCGATTAATGACAGAGTCAGATATCTTCAGGGTTGCTATT 2280
Qy 761 AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780
Db 2281 GCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACTTATATG 2340
Qy 781 GlyHisProAspGlnAsnGluGlnGlyTyrTyrIleGlySerValAlaMetGlnAlaGlu 800
Db 2341 GGTCAACCTTGCACAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAA 2400
Qy 801 LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
Db 2401 AGTTCCCTCTGACCAAAATCGTTTACTGCTCTTACATGGTTTCTCGGATGAGATGTC 2460
Qy 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
Db 2461 CATTTTGCATACACAGTATATTAATGAGTTTTTTAGTGGGGCTGGAAAGCCATATGAT 2520
Qy 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
Db 2521 TTACAGATCTATCTCTCAGGAGACACACAGTAAAGTTTCTGATCGGAGAACATTAT 2580
Qy 861 GluLeuHisLeuLeuHisTyrIleGlnGluAsnLeuGlySerArgIleAlaLeuLys 880
Db 2581 GAACTCATCTTTTGGCACTACCTTCAAGAAACCTTGGATCAGGTATTTGCTGCTCTAAAA 2640
Qy 881 ValIle 882
Db 2641 GTGATA 2646
```

RESULT 2

```
US-11-151-601-19
; Sequence 19, Application US/11151601
; Publication No. US20060003413A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Olandt, Peter J.
; APPLICANT: Kapeller-Libermann, Rosana
```

```
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Williamson, Mark
; APPLICANT: Weich, Nadine
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASE, PHOSPHATASE,
; FILE REFERENCE: AND PROTEASE FAMILY MEMBERS AND USES THEREOF
; CURRENT APPLICATION NUMBER: US/11/151,601
; CURRENT FILING DATE: 2005-06-13
; PRIOR APPLICATION NUMBER: US 10/170,789
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: US 09/797,039
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: PCT/US01/06525
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/186,061
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 09/882,166
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: PCT/US01/19269
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/212,078
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/934,406
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: PCT/US01/26052
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/226,740
; PRIOR FILING DATE: 2000-08-21
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 3143
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (229) ... (2874)
US-11-151-601-19
```

```
Alignment Scores:
Pred. No.: 0 Length: 3143
Score: 4700.00 Matches: 882
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 14 Gaps: 0

US-10-825-632-1 (1-882) x US-11-151-601-19 (1-3143)
Qy 1 MetAlaAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
Db 229 ATGCGACGACCAATGGAACACAGACGCTGGGTGTGAGATATTTGAACTGCGGACTGT 288
Qy 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
Db 289 GAGGAGAAATATTGAATCACAGGATCGGCTAAATTTGGAGCCCTTTTATGTTGAGCGGTAT 348
Qy 41 SerTyrSerGlnLeuLysLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
Db 349 TCCTGGAGTCAGCTTAAAAAGCTGTGCGGATACACAGAAATATCATGGCTACATGATG 408
Qy 61 AlalaAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
Db 409 GCTAAGGCACCAATGATTTTCATGTTTGTGAAGAGAAATGATCCAGATGGACCTCATCA 468
Qy 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
Db 469 GACAGAAATCTATTACCTTCCCATGCTGTGTGAGAGACAGAGAAATATACACTGTTTATTCT 528
Qy 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTyrLysProLeuLeu 120
Db 529 GAAATTTCCCAAACTATCAATAGAGCAGCAGTCTTAAATGCTCTCTTGGAGCCCTCTTTTG 588
```

Qy	121	AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg	140	Qy	481	GlyLeuProAlaProSerAspPheLeuCysProIleLeuGluGluLeuAlaIleThrSer	500
Db	589	GATCTTTTTCAGGCAACACTGGACTATGGAAATGATATTCGAGAAAGAACTATTAAGA	648	Db	1669	GGGCTGGCTGCTCCAGTGATTCAGTGCTCCATCAAGAGGAGATAGCAATACCAGT	1728
Qy	141	GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly	160	Qy	501	GlyGluThrGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg	520
Db	649	GAAGAAGAAACGATTTGGAACAGTCGGAAATTCCTTACGATATTAACAAGGAAGTGGG	708	Db	1729	GGTGAAATGGGAAGTTCTTGGCCGCGCATGGATCTATATCCAGTTGATGAGTCAGAGG	1788
Qy	161	ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly	180	Qy	521	LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer	540
Db	709	ACATTTCTGTTTCAAGCGGTAGTGGAAATTTATACGTAAAGATGGAGGGCCACAAGGA	768	Db	1789	CTGGTATATTTGAAGGACCAAGACTCCCTTTTAGAGCATCCTCTACGTAGTCACT	1848
Qy	181	PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet	200	Qy	541	TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys	560
Db	769	TTTACGCCAACACCTTTAAGGCCCAATCTAGTGGAAACTAGTTGTGCCCAACATACGGATG	828	Db	1849	TACGTAATACTCGAGAGGTCACAAGGCTGACACCGTGGCTACTCACATTTCTTCTCTC	1908
Qy	201	AspProLysLeuCysProAlaAspProAspTyrIleAlaPheIleHisSerAsnAspIle	220	Qy	561	IleSerGlnHisCysAspPheIleSerLysTyrSerAsnGlnLysAsnProHisCys	580
Db	829	GATCCAAATATATGCCCTGCTGATCCAGACTGGATTCCTTTATACATAGCAACGATATT	888	Db	1909	ATCAGTCAGCACTGTGACTCTCTTTATTAAGTAAGTATAGTAACCAAGAAATCCACACTGT	1968
Qy	221	TrpIleSerAsnIleValThrArgGluArgArgLeuThrTyrValHisAsnGluLeu	240	Qy	581	ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe	600
Db	889	TGGATATCTAACATCGTAACCGAGAGAAAGAGAGACTCACTTATGTGCACATGAGCTA	948	Db	1969	GTGTCCCTTTACAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAAAAACAAAGGAATTT	2028
Qy	241	AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu	260	Qy	601	TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePhe	620
Db	949	GCCAACTGGGAAGAGATGCCAGATCAGCTGGAGTCGCTACCTTGTCTCCAGAGAA	1008	Db	2029	TGGGCCACCATTTGGATTTCAGCAGGTCTCTTCTGACTATATCTCTCCAGAAATTTTC	2088
Qy	261	PheAspArgTyrSerGlyTyrTrpTyrProLysAlaGluThrThrProSerGlyGly	280	Qy	621	SerPheGluSerThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu	640
Db	1009	TTTGTAGATATATCTGGCTATTTGGTGGTGTCCAAAGCTGAACAACTCCCAAGTGGTGT	1068	Db	2089	TCTTTTGAAGTACTACTGGATTTACATTTGATGGATGCTCTACAAGCTCATGATCTA	2148
Qy	281	LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal	300	Qy	641	GlnProGlyLysGlyTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu	660
Db	1069	AAAAATCTTAGAAATCTATATGAAGAAATGATGAATCTGAGGTGGAATTAATCATGTT	1128	Db	2149	CAGCTCGGAAGAAATATCTTACTGTGCTGTTCATATATGTTGGTCTCCAGGTGCAGTTG	2208
Qy	301	ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr	320	Qy	661	ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly	680
Db	1129	ACATCCCTATGTTGAAACCAAGAGGGGCAGATTTCATTCGTTATCTCTAAAAACAGGTACA	1188	Db	2209	GTGAATAATCGGTTTAAAGGAGTCAGTATTTCCGTTGAATACCTTAGGCTCTTAGGT	2268
Qy	321	AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle	340	Qy	681	TyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly	700
Db	1189	GCAAAATCCTTAAAGTCACATTTTAAAGATGTCAGMAATATATGATTGATCTGAAGGAGATC	1248	Db	2269	TATGTGGTTGATGATAGACAACACAGGGGATCCTGTCCAGGAGGCTTAATTTGAAGGC	2328
Qy	341	IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu	360	Qy	701	AlaPheLysTyrLysMetGlyGlnIleGluLeuAspAspGlnValGluGlyLeuGlnTyr	720
Db	1249	ATAGATGTCATAGATAAGGAACTAATTCACCTTTTCAGATTCTATTTGAAGGAGTTGAA	1308	Db	2329	GCCTTTAAATATAAAATCGGTCAAATAGAAATTCAGATCAGGTGGAAGGACTCCAATAT	2388
Qy	361	TyrIleAlaArgAlaGlyThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp	380	Qy	721	LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyr	740
Db	1309	TATATTGCCAGAGCTGGATGGACTCTCGAGGAAATATGCTTGGTCCATCTCTACTAGAT	1368	Db	2389	CTAGCTTCTCGATATGATTTTCATTGACTTAGATCTGTGGGCATCCACGGCTGGTCTAT	2448
Qy	381	ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu	400	Qy	741	GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle	760
Db	1369	CGCTCCAGACTCGCTCGAGATGTTGATCTCACCTGATATTTATATCCAGTAGAA	1428	Db	2449	GGAGGATACCTCTCCCTGATGGCATTAATGCAGAGGTCAGATATCTTCAGGGTTGCTATT	2508
Qy	401	AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu	420	Qy	761	AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet	780
Db	1429	GATGATGTTATGGAAGGCAGAGACTCAATGAGTCAGTGGCTGATCTGTGAGCCACCTA	1488	Db	2509	GCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACAGGAACGTTATATG	2568
Qy	421	IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe	440	Qy	781	GlyHisProAspGlnAsnGluGlnGlyTyrTyrTyrLeuGlySerValAlaMetGlnAlaGlu	800
Db	1489	ATTATCTATGAAGAACACACAGACATCTGGATAAATATCCATGACATCTTTCAGTGT	1548	Db	2569	GSTCACCTTGACCAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGACAGAA	2628
Qy	441	ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe	460	Qy	801	LysPheProSerGlnProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal	820
Db	1549	CCCAAGTCACGAAGAGGAAATGAGTTATTTTTCCTCTGAATGCAAAAACAGGTTTC	1608	Db	2629	AAGTTCCTCTGACCAAAATCGTTTACTGTCTTACATGGTTCTTCGGATGAGAATGTC	2688
Qy	461	ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly	480	Qy	821	HisPheAlaHisThrSerIleLeuSerPheLeuValArgAlaGlyLysProTyrAsp	840
Db	1609	CGTCAATTATACAAAATTAACATCTATTTTAAAGGAAAGCAAAATATAAACATCCAGTGT	1668	Db	2689	CATTTGCACTACCAAGTATATTTACTGAGTTTTTTAGTGAGGGCTGGAAAGCCATATGAT	2748
Db				Qy	841	LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr	860

Db 2749 TTACAGATCTATCTCTCAGGAGACACAGCATTAAGAGTTCTCGAATCGGAGAACATTAT 2808
 Qy 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLys 880
 Db 2809 GNACTGCATCTTTTGCACTACTCTCAAGAAACCTTGGATCAGTATTGCTCTCTAAA 2868
 Qy 881 Vallie 882
 Db 2869 GTGATA 2874

RESULT 3

US-11-079-463-1186
 ; Sequence 1186, Application US/11079463
 ; Publication No. US20060073161A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary L. Breton
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTEROIDES FR
 ; FILE REFERENCE: PATH00-03DIV2
 ; CURRENT APPLICATION NUMBER: US/11/079,463
 ; PRIORITY FILING DATE: 2005-03-14
 ; PRIOR APPLICATION NUMBER: US 60/128,705
 ; PRIOR FILING DATE: 1999-04-09
 ; PRIOR APPLICATION NUMBER: US 09/540,209
 ; PRIOR FILING DATE: 2000-04-04
 ; NUMBER OF SEQ ID NOS: 10444
 ; SEQ ID NO 1186
 ; LENGTH: 2238
 ; TYPE: DNA
 ; ORGANISM: B. fragilis
 US-11-079-463-1186

Alignment Scores:
 Pred. No.: 1,76e-59 Length: 2238
 Score: 651.00 Matches: 199
 Percent Similarity: 44.3% Conservative: 124
 Best Local Similarity: 27.3% Mismatches: 244
 Query Match: 13.9% Indels: 162
 DB: 11 Gaps: 29

US-10-825-632-1 (1-882) x US-11-079-463-1186 (1-2238)

Qy 173 VallysAspGlyGlyProGlnGlyPheThrGlnGlnProLeuArgProAsnLeuValGlu 192
 Db 448 CTATCGAGCGTGTGTCGCGAG-----468
 Qy 193 ThrSerCysProAsnIleArgMetAspProLysLeuCysProAlaAspProAspTrpIle 212
 Db 469 -----CAAGCCCGGTATTCTCTCCG---GACGGAACCTGTGTG 504
 Qy 213 AlaPheIleHisSerAsnAspIleTrpIleSerAsnIleValThrArgGluArgArg 232
 Db 505 GCGTTTGTGCGGTGTAATAACACTCTCTCGTA-----AAG 540
 Qy 233 LeuThrTyrValHisAsnGluLeuAlaAsnMetGluAspAlaArg-----248
 Db 541 CTTTGTATGGCAACAGCGAA---TCGCAAGTGAAGTGAACCGCAAGTTGAACGTGTA 597
 Qy 249 SerAlaGlyValAlaThrPheValLeuGlnGlu-----PheAspArgTyrSerGly 266
 Db 598 CTGAACGGTATTCCCGACTGGGTATACGAAGAGTTCGGTTTCAACCGTCCCTGGAA 657
 Qy 267 TyrTrpTrpCysProGlyAlaGluThrThrProSerGlyGlyLysIleLeuArgIleLeu 286
 Db 658 TTC-----AATGCCGATAATACC-----ATGCTGGCC 684
 Qy 287 TyrGluGluAsnAspGluSerGluValGluIleIleHisValThrSerProMetLeu---305
 Db 685 TATGTTGTTTCGACGAATCGAGGTTCAC-----TCATACACTTTCCTCCCTGTTGCA 738
 Qy 306 ---GluThrArgArgAlaAsp-----SerPheArgTyr 315
 Db 1739 GGTGAAGCACCAGCGTTATGATGTCACATCGAGGATTATCCGGGAGAAATACACTTACAAATAT 7398
 Qy 316 ProLysThrGlyThrAlaAsnProLysVal-----ThrPhe-----327
 Db 799 CCAAGAGCAGGTATCCCAACTTCCCAAGGTGTGACATACATCGTTTCGACATCAAAATCGAAA 858
 Qy 328 -----LysMetSerGluIleMetIleAspAlaGluGlyArgIleIleIleAspValIleAsp 345
 Db 859 GTGACCCGTCAGGTGAAGCTGCCGATAGACCCCGCGA-----897
 Qy 346 LysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGluTyrIleAlaArgAla 365
 Db 898 -----TATATCCCGCGCATC 912
 Qy 366 GlyTrpTrpProGluGlyLysTyrAlaTrpSerIleLeuLeuAspArgSerGlnThrArg 385
 Db 913 CGTTTCACTCAGGATCCCAACAACTGCGCATCATGACACTGAAACCGTCACCAAGAACCGC 972
 Qy 386 LeuGlnIleValLeuIleSerProGluLeuPheIleProValGluAspValMetGlu 405
 Db 973 TTCGACATGATATTTTGGCGATCCT-----996
 Qy 406 ArgGlnArgLeuIleGluSerValProAspSerValThrProLeuIleIleTyrGluGlu 425
 Db 997 CGCAGCAGAGTGTGCAAACTGCGCCCTACGCGACGAATCTCTTATTATCATCAACGAA--1053
 Qy 426 ThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPheProGlnSerHisGlu 445
 Db 1054 -----AATGATTCGATAAATTCAGTTCTTCTATCCGAA-----1086
 Qy 446 GluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPheArgHisLeuTyrLys 465
 Db 1087 -----TATTTTCACTCTTTGTAGCGAT---AAGACGGATATCTCTACTTGTATC---1131
 Qy 466 IleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGlyGlyLeuProAlaPro 485
 Db 1132 -----TGGTATAGCATGAACGTAACCTG-----1155
 Qy 486 SerAspPheLysCysProIleLysGluIleAlaIleThrSerGlyGluTrpGluVal 505
 Db 1156 -----ATCAACAA-----GTGACACGCGGTAACTGAAGTA 1188
 Qy 506 LeuGlyArgHisGlySerAsnIleGlnValAspGluValArgLeuValTyrPheGlu 525
 Db 1189 AAAAACTTTATTCGGATGGAATCCGATACCAACAG-----TTTATTATACAC 1236
 Qy 526 GlyThrLysAspSerProLeuGluHisIleLeuTyrValSerTyrValAsnProGly 545
 Db 1237 AGCAATGAAGAAGCCCGATGCGTCAGCGGTATACAGATAGAC-----CGTAAGGCG 1290
 Qy 546 GluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCysIleSerGlnHisCys 565
 Db 1291 AAGAAGTGAATCTGAGCAATCAGCGGGAACCAACAGTCCCATCTTCAGCAGCTCGATG 1350
 Qy 566 AspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSerLeuTyrLys 585
 Db 1351 AATATATTATGAACAAGTTTACCAGCTCGATATCTCGATGCTGATTTACCTTG-----1404
 Qy 586 LeuSerSerProGluAspAspProThrCysLysThrLysGluPheTrpAlaThrIleLeu 605
 Db 1405 -----AATGACCAACAGATGAGTCTTGAAGACTCTCGTAACAAATGAT 1449
 Qy 606 AspSerAlaGlyProLeuProAspTyrThrProPro-----GluIlePheSerPheGlu 623
 Db 1450 AAATGAACAGAACTGCGCGAATATGCCATACCGCAAAAGAAATTTCTTCACGTTCAA 1509
 Qy 624 SerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 643
 Db 1510 ACAACAGAGAGGAGTTCGATCTGAACGCTGGATGATGAACCGCTCAATTTTCGATCCTGCC 1569
 Qy 644 LysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsnAsn 663
 Db 1570 AAACGTTATCCGGTACTGATGTTCCAGTATAGCGGTCCGGTTCGCAACAGGTTCTGAC 1629

```
QY 664 ArgPheLysGlyVallystYrPheArgLeuAenThrLeuAlaserLeuGlyTyrValVal 683
DB : : : : :
DB 1630 AAATGG--GGAATCAGTTGG-----GAAACCTCATGCGCAGCTCGGTACGTGGTA 1680
QY 684 ValValIleAspAenArgGlySerCysHisArgGlyLeuLysPheGluGlyValaPheLys 703
DB : : : : :
DB 1681 GCTTGTGTAGATGTCGCGGCACAGGTGGCGGTGGCAGTGAATTCAGAAAATGCCACTTAC 1740
QY 704 TyrLysMetGlyGlnIleLeuAspAenValGlnValGlnGlyLeuGlnTyrLeuAlaser 723
DB : : : : :
DB 1741 CTGAACCTGGGTGTAAGAAAGAGCTAAAGACAGGTGGAAGCTGCCAAATATCTGGGTGGA 1800
QY 724 ArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyrGlyGlyTyr 743
DB : : : : :
DB 1801 ---CTGCCTTATGTGGCAAGGACGTATTTGGTATCTGGGATGGAGTTTCGGCGGATAT 1857
QY 744 LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAla 763
DB : : : : :
DB 1858 ATGACCATCATGAGTATGAGGAGGTACACCGTGTGTAAAGCCGGAGTTGCTGTGCC 1917
QY 764 ProValThrLeuTyrIlePheTyrAspThrGlyTyrThrGluArgTyrMetGlyHisPro 783
DB : : : : :
DB 1918 GCACCTACAGACTCGGAATATTACGATACAGTATATACCGAACGCTTTATGCGCAGCG 1977
QY 784 AspGlnAenGlnGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlyLysPhePro 803
DB : : : : :
DB 1978 AAAGAAAATGCCGAAGGCTATAAAGCAGCTTCAGCATTCAGCGGTGCAGACCAACCTG-- 2034
QY 804 SerGluProAenArgLeuLeuLeuHisGlyPheLeuAspGluAenValHisPheAla 823
DB : : : : :
DB 2035 -----CATGTAACCTGCTCTTGTATACCGGTATGCGCAGATGATGTAATGTTCACTTCCAG 2088
QY 824 HisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAspLeuGlnIle 843
DB : : : : :
DB 2089 AACTGTACAGAATATGCGAGACACCTGTGTACAACTCGSAAAACAGTTTCGATATGCAAGTA 2148
QY 844 TyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyrGluLeuHis 863
DB : : : : :
DB 2149 TACACCAACCGAATCATAGCATCTATGTTGGAATACCCGTAAACCCTTGTATACGAAG 2208
QY 864 LeuLeuHisTyrLeuGlnGluAenLeu 872
DB : : : : :
DB 2209 CTGACGAACCTTCTCCGGATATATTA 2235

RESULT 4
US-11-079-463-2536
; Sequence 2536, Application US/11079463
; Publication No. US20060073161A1
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTEROIDES FRA
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: PATH00-03DIV2
; CURRENT APPLICATION NUMBER: US/11/079,463
; CURRENT FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/128,705
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: US 09/540,209
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO: 2536
; LENGTH: 1884
; TYPE: DNA
; ORGANISM: B. fragilis
; US-11-079-463-2536

Alignment Scores:
Pred. No.: 9,718-49 Length: 1884
Score: 551.00 Matches: 188
Percent Similarity: 44.1% Conservative: 120
Best Local Similarity: 25.7% Mismatches: 249
Query Match: 11.7% Indels: 116
```

```
Db 1000 TACCTGATTGACCAATACCACTCACCAGCTGTTCGAGAGATATCAATATCATCGATGG 1059
Qy 587 SerSerProGluLeuAspProThrCysLysThrLysGluPheThrPalaThrLeuLeuAsp 606
Db 1060 CAAGTGGTAAAGC-----GTCAATCTGCTCACA 1089
Qy 607 SerAlaGlyProLeuProAspTyrThrProPro-----GluLeuPheSerPheGluSer 624
Db 1090 GCAGCCGATCCTTTACCCGGCTACAAGATGCGGGTATCGAGACAGAGTACCATCAAGCA 1149
Qy 625 ThrThrGlyPheThr---LeuTyrGlyMetLeuTyrLysProHisAspLeuGlnProGly 643
Db 1150 GCAGATGGCAAGACGACTTATATTACCGATTATATAAAACCGGCCGATTTTCGATCCGAAC 1209
Qy 644 LysLysTyrProThrValLeuPheLeuTyrGlyProGlnValGlnLeuValAsnAsn 663
Db 1210 AAGAAATACCCGGCTATCGTATACGATACGCGGTCTCTCAGCAGCACTGTGTCCACAC 1269
Qy 664 ArgPheLys---GlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGlyTyrVal 682
Db 1270 GGCTGGCAGAACGGTGCACGAGCTGGGATATC---TATATGGCCAAACAAAGGTTATATC 1326
Qy 683 ValValValLeuAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPhe 702
Db 1327 ATGTTTACCGTAGACCGAGCTGGAAGCAGCAATCGGGACTCGATTTTGAATGTGTACT 1386
Qy 703 LysTyrLysMetGlyGlnLeuGluLeuAspGlnValGluGlyLeuGlnTyrLeuAla 722
Db 1387 TTCGGCCAGTTGGGATTCGAAGAAGAGACAGGAGGGAAGGAACCGCAATCTCTGAAA 1446
Qy 723 SerArgTyrAspPheLeuAspLeuArgValGlyLeuHisGlyTyrSerTyrGlyGly 742
Db 1447 AGC---CTCCCTACGTCGTGGAACCGTATCGGAGTACACGGCTGGAGCTTTGGCGGT 1503
Qy 743 TyrLeuSerLeuMetAlaLeuMetGlnArgSerAspLeuPheArgValAlaLeuAlaGly 762
Db 1504 CACATGACCACTGCCCTTCTCTTCGCTATCCGAGATATTTAAAGTGGCGGTGGCGGC 1563
Qy 763 AlaProValThrLeuThrPheTyrAspThrGlyTyrThrGluArgTyrMetGlyHis 782
Db 1564 GGTCCCGCTCATTCAGCTGGGGGTATTACGAAGTGTGTACGAGAACGCTATATGGATACC 1623
Qy 783 ProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGluLysPhe 802
Db 1624 CCGCAAGCAATCGAAAGGATATAAAGAGTGTAACTGAAACCTGGCAGGCAACCTG 1683
Qy 803 ProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnValHisPhe 822
Db 1684 -----AAAGGACATCTGATGATCATCCAGATGACACGACGACACCTGGCGTCCG 1734
Qy 823 AlaHisThrSerLeuLeuSerPheLeuValArgAlaGlyLysProTyrAspLeuGln 842
Db 1735 CAGCACACCTCTCGTTTATGAAGCGTGCATCGATGCCCGCAGCATATCGGACCTCTTC 1794
Qy 843 IleTyrProGlnGluArgHisSerLeuArgValProGluSerGlyGlu-----His 859
Db 1795 ATCTACCCGTCGCAAGCAACAATGTG-----TCGGCGCGGACCGGTGTACAT 1842
Qy 860 TyrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeu 872
Db 1843 CTGCATGAAAGATAACCCGTTATTTTCGAAGACTATTTA 1881
```

RESULT 5

US-11-208-288-3

; Sequence 3, Application US/11208288

; Publication No. US20060051366A1

; GENERAL INFORMATION:

; APPLICANT: CHANG, Chiwen

; TITLE OF INVENTION: USE OF SOLUBLE CD26 AS INHIBITOR OF

; TITLE OF INVENTION: ANGIOGENESIS AND INFLAMMATION

; FILE REFERENCE: 39533-0001

; CURRENT APPLICATION NUMBER: US/11/208,288

; CURRENT FILING DATE: 2005-08-18

; PRIOR FILING DATE: 2004-08-26

; NUMBER OF SEQ ID NOS: 8

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 3

; LENGTH: 2217

; TYPE: DNA

; ORGANISM: Homo sapiens

; US-11-208-288-3

Alignment Scores: 2,96e-46 Length: 2217
Pred. No.: 529.00 Matches: 186
Score: 529.00 Conservative: 109
Percent Similarity: 42.0% Mismatches: 280
Best Local Similarity: 26.5% Indels: 128
Query Match: 11.3% Gaps: 32
DB: 11

US-10-825-632-1 (1-882) x US-11-208-288-3 (1-2217)

```
Qy 206 ProAlaAspProAspTrpIle-----AlaPheIleHis 216
Db 361 CCAACACACACAGTGGGTGCATGTCACAGTGGGTGCATATAAATGGCATATGTTGG 420
Qy 217 SerAsnAspIleTrpIleSerAsnIleValThrArgGluGluArgArgLeuThrTyrVal 236
Db 421 AACAAATGACATTTATGTTAAATTTGAACCAATTTTACCAAGTTTACAGAAATCAGATGACG 480
Qy 237 HisAsnGluLeuAlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheVal 256
Db 481 GGG-----AAAGAAGATATAATATATAATGAATAAAGTACTGGGTT 522
Qy 257 LeuGlnGluGlu---PheAspArgTyrSerGlyTyrTrpTyrCysProLysAlaGluThr 275
Db 523 TATGAAGAGGAGTCTTTCAGTGCCTACTGCTGTGTGTGTGTCTCTCCAAACGGCAGCTTT 582
Qy 276 ThrProSerGlyGlyLysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluVal 295
Db 583 -----TTAGCATATGCCCAATTTAAACGACACACAGAAGTC 615
Qy 296 GluIleIleHis-----ValThrSerProMetLeuGluThrArgArgAlaAspSer 312
Db 616 CCACATTATGAATPACTCTTCTACTGATGAGTCACGTCAGTACCCAAAGACTGTACGG 675
Qy 313 PheArgTyrProLysThrGlyThrAlaAsnProLysValThrPheLysMetSerGluIle 332
Db 676 GTTCCATATCCAAAGCGCAGGAGCTGTGAATCCCACTGTAAAGTTCTTTGTT-----GTA 729
Qy 333 MetIleAspAlaGluGlyArgIleIleAspValIleAspLysGluLeuIleGlnProPhe 352
Db 730 AATACAGACTCTCTCAGCTCAGTCACCAATGCAACTTCCATACAAATCACTGCTCTGCT 789
Qy 353 GluIleLeuPheGluGlyValGluTyrIleAlaArgAlaGlyTyrThrProGluGlyLys 372
Db 790 TCTATGTTGATA---GGGGATCACTACTGTGTGTGTGTGACATGGCACTGGCA----- 834
Qy 373 TyrAlaTrpSerIleLeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSer 392
Db 835 -----ACACAGAAAGAAATTTCTTTCAGTGGCTC----- 864
Qy 393 ProGluLeuPheIleProValGluAspValMetGluArgGlnArgLeuIleGluSer 412
Db 865 -----AGGAGGATTCAGAAC 879
Qy 413 ValProAspSerValThrProLeuIleIleTyrGluGluThrThrAspIleTrpIleAsn 432
Db 880 TAT-----TCGGTTCATGATATTTGTGACTATGATGAATCCAGTGAAGATGG---AAC 930
Qy 433 IleHisAspIlePheHisValPheProGlnSerHisGluGluGluIleGluPheIlePhe 452
Db 931 TGC-----TTAGTGGCAGCGGCAACACATTGAATGAGTACTACTGCTGGTGG 975
```

```
QY 453 AlaserGluCysLysThrGlyPheArgHis-----LeuTyrLys 465
Db 976 GTTGAAGATTAGGCTTACAGACCTCATTTTACCTTGATGGTAATAGCTTTCACAG 1035
QY 466 lIeThrSerIleLeuLysGluSerLysTyrLysArgSerSerGlyGlyLeuProAlaPro 485
Db 1036 ATCATCAGC-----AATGAAGAAGGTTACAGACACATT----- 1068
QY 486 SerAspPheLysCysProIleLysGluIleAlaIleThrSerGlyGluTrpGluVal 505
Db 1069 TCGTATTTCCAAATAGATAAAGAGAGCTGCACATTATTACAAAAGCCACCTGGGAATC 1128
QY 506 LeuGlyArgHisGlySerAsnIleGlnValAspGluValArgGluValTyrPheGlu 525
Db 1129 ATCGGG-----ATAGAAGCTCTAACCAAGTATTATCTATCTACATTAGT 1173
QY 526 GlyThr-----LysAspSerProLeuGluHisLeuTyrValValSerTyrValAsnPro 544
Db 1174 AATGAATATAAAGGATCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1233
QY 545 GlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysLysLeuSer--Gln 563
Db 1234 ACAAAAGTGACATGCTC-----AGTTGAGCTGATCCGGA 1272
QY 564 HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHis-----Cys 580
Db 1273 AGGTGTCAGTACTATTCTGTCTCATTAGTAAGAGCGGAGGAGGAGGAGGAGGAGGAGG 1332
QY 581 -----ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLys 596
Db 1333 TCCGGTCTCTGCTCCCTCTATCTCTACAGCAGCGTGAATATAAGCGGTGAGA 1392
QY 597 ThrLysGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrPro 616
Db 1393 GTCCTGGAAGAC---AATTGAGCTTTGGATAAAATG-----CTGCAGAAATGCCAGATG 1443
QY 617 ProGlu-----IlePheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMet 633
Db 1444 CCTCCAAAAGAACTGGACTTCATTATTTTGAATCAAAACAAATTT---TGGTATCAGATG 1500
QY 634 LeuTyrLysProHisAspLeuGlnProGlyLysLysTyrProThrValLeuPheIleTyr 653
Db 1501 ATCTGCTCTCTCAT---TTTGATAAATCCAGAAATATCTCTACTATTAGATGTAT 1557
QY 654 GlyGlyProGlnValGlnLeuValAsnAsnArgPheLysGlyValLysTyrPheArgLeu 673
Db 1558 GCAGGCCCATGAGTACAA-----AAAGCAGACACTGCTTCAGACTG 1599
QY 674 Asn-----ThrLeuAlaSerLeuGlyTyrValValValVal-----IleAspAsnArg 689
Db 1600 AACTGGGCCACTTACCTTCAAGCACAGAAACATATTAGTAGCTAGCTTTGATGCAGA 1659
QY 690 GlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLysTyrLysMetGlyGlnIle 709
Db 1660 GGAAGTGGTTTACCAGAGATGAAGATCATGTCATGCAATCAACAGAGACTGGGAACATTT 1719
QY 710 GluIleAspAspGlnValGluGlyLeuGlnTyrLeuAlaSerArgTyrAspPheIleAsp 729
Db 1720 GAAGTTGAAGATCAAAATGAAGCAGCAGCAAAATTT---TCAAAAATGGGATTTGGGAC 1776
QY 730 LeuAspArgValGlyIleHisGlyTrpSerTyrGlyLysTyrLeuSerLeuMetAlaLeu 749
Db 1777 AACAAAGAAATTCGAATTTGGGCTGCTCATATGAGGAGTACGTAACCTCAATGGTCTCG 1836
QY 750 MetGlnArgSerAspIlePheArgValAlaIleAlaGlyAlaProValThrLeuTrpIle 769
Db 1837 GGATCGGAAGTGGCGTGTCAAGTGTGGAATAGCGGTGGCGCTGTATCCCGTGGGAG 1896
QY 770 PheTyrAspThrGlyTyrThrGluArgTyrMetGly-----HisProAspGlnAsnGlu 787
Db 1897 TACTATGACTGATGTACAGAACGTTTACATGGGTCTCCCACTCCAGAGACACCTT 1956
QY 788 GlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGluLysPheProSerGluProAsn 807
```

```
Db 1957 GACCATTACAGAAATTCACAGTCATGAGCAGAGCTGAAAAATTTT-----AAACAAGTT 2010
QY 808 ArgLeuLeuLeuLeuHisGlyPheLeuAspGluAsnValHisPheAlaHisThrSerIle 827
Db 2011 GAGTACCTCTTATTCATGGAACAGCAGATGATAACGTTTCACTTTTCAGCAGTCAGCTCAG 2070
QY 828 LeuLeuSerPheLeuValArgAlaGlyLysProTyrAspLeuGlnIleTyrProGlnGlu 847
Db 2071 ATCTCCAAAGCCCTGGTCGATGTTGGATGGATTTCCAGGCAATGTGGTATATCTGATGAA 2130
QY 848 ArgHisSerIleArgValProGluSerGlyGluHisTyrGluLeuHisLeuLeuHisTyr 867
Db 2131 GACCATTGAATAGTAGTAGCAGACACCAACATATATATATATATATATATATATATATAT 2190
QY 868 LeuGlnGlu 870
Db 2191 ATAAACAA 2199

RESULT 6
US-11-208-288-1
; Sequence 1, Application US/11208288
; Publication No. US20060051366A1
; GENERAL INFORMATION:
; APPLICANT: CHANG, Chiwen
; TITLE OF INVENTION: USE OF SOLUBLE CD26 AS INHIBITOR OF
; TITLE OF INVENTION: ANGIOGENESIS AND INFLAMMATION
; FILE REFERENCE: 39533-0001
; CURRENT APPLICATION NUMBER: US/11/208,288
; PRIOR FILING DATE: 2005-08-18
; PRIOR FILING DATE: 2004-08-26
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3332
; TYPE: DNA
; ORGANISM: homo sapiens
US-11-208-288-1

Alignment Scores:
Pred. No.: 5,23e-46 Length: 3332
Score: 529.00 Matches: 186
Percent Similarity: 42.0% Conservative: 109
Best Local Similarity: 26.5% Mismatches: 280
Query Match: 11.3% Indels: 128
DB: 11 Gaps: 32

US-10-825-632-1 (1-882) x US-11-208-288-1 (1-3332)
QY 206 ProAlaAspProAspTrpIle-----AlaPheIleHis 216
Db 445 CCAACACACACACAGTGGGTGCACATGTCACAGTGGGTGCATAAATTTGGCATATGTTGG 504
QY 217 SerAsnAspIleTrpIleSerAsnIleValThrArgGluArgLeuThrTyrVal 236
Db 505 AACAAATGACATTTATGTTAAATTCACAAATTTTACCAAGTTACAGAAATCATCATGAGC 564
QY 237 HisAsnGluLeuAlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheVal 256
Db 565 GGG-----AAAGAGATATATATATATATATATATATATATATATATATATATATAT 606
QY 257 LeuGlnGlu-----PheAspArgTyrSerGlyTyrTyrTrpCysProLysAlaGluThr 275
Db 607 TATGAAGAGAGAACTTTCAGTCCCTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 666
QY 276 ThrProSerGlyGlyLysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluVal 295
Db 667 -----TTAGCATATGCCCAATTTTACGACACAGACAGATC 699
QY 296 GluIleIleHis-----ValThrSerProMetLeuGluThrArgArgAlaAspSer 312
Db 700 CCACTTATTGATACTCTTCTACTGATGAGTCACAGTACCCCAAGACTGTACGG 759
```


Qy 313 PheArgTyrProLysThrGlyThrAlaAsnProLysValThrPheLysMetSerGluIle 332
Db 760 GTTCCATATCAAAAGCAGCAGCTGTGAATCACTGTAAGTTCTTTGTT-----GTA 813
Qy 333 MetIleAspAlaGluIleArgLysValIleAspLysGluLeuIleGlnProPhe 352
Db 814 AATACAGACTCTCTCAGCTCAGTCACCAATCCCACTTCCATACAAATCAGCTCCTGCT 873
Qy 353 GluIleLeuPheGluGlyValGluTyrIleAlaArgAlaGlyTyrThrProGluGlyLys 372
Db 874 TCTAATGTTGATA---GGGATCACTACTTGTGTGTGATGACATGGCA----- 918
Qy 373 TyrAlaTrpSerIleLeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSer 392
Db 919 -----ACACAAGAAAGAAATTTCTTTCAGTGGCTC--- 948
Qy 393 ProGluLeuPheIleProValGluAspAspValMetGluArgGlnArgLeuIleGluSer 412
Db 949 -----AGGAGGATTCAGAAC 963
Qy 413 ValProAspSerValThrProLeuIleIleTyrGluGluThrThrAspIleTyrIleAsn 432
Db 964 TAT-----TCGGTCATGATATTTGTGATGATGATATCCAGTGGAGATGG---AAC 1014
Qy 433 IleHisAspIlePheHisValPheProGlnSerHisGluGluIleGluPheIlePhe 452
Db 1015 TGC-----TTAGTGGCAGCGCAACACATTGAATGAGTACTACTGGCTGG 1059
Qy 453 AlaSerGluCysLysThrGlyPheArgHis-----LeuTyrLys 465
Db 1060 GTTGGAGAGATTAGGCTTCAGAACCTCATTTTACCCTTGATGGTAATAGCTTCTACAG 1119
Qy 466 IleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGlyGlyLeuProAlaPro 485
Db 1120 ATCATCAGC-----AATGAAGAAGGTTACACACACATT----- 1152
Qy 486 SerAspPheLysCysProIleLysGluGluIleAlaIleThrSerGlyLutrpGluVal 505
Db 1153 TGCTATTTCCTCAATAGATAAAGAGTCACTATTTTCAAAAGGACCTGGGAGTC 1212
Qy 506 LeuGlyArgHisGlySerAniIleGlnValAspGluValArgArgLeuValTyrPheGlu 525
Db 1213 ATCGGG-----ATAGAAGCTCTAACCCAGTGAATTATCTATCTACTACTATTAGT 1257
Qy 526 GlyThr---LysAspSerProLeuGluHisLeuTyrValValSerTyrValAsnPro 544
Db 1258 AATGAATATAAAGGAATGCCAGGAGGAAGGAATCTTTATAAAATCCAACTTATTGACTAT 1317
Qy 545 GlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCysIleSer---Gln 563
Db 1318 ACNAAAGTGACATGCCTC-----AGTTGTGAGCTGAATCCGGAA 1356
Qy 564 HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHis-----Cys 580
Db 1357 AGGTGTCAGTACTATTCTGTGTCATTTCAGTAAGAGCGCAAGTATTATCAGCTGAGATGT 1416
Qy 581 -----ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLys 596
Db 1417 TCCGTCCTGCTGCTGCCCTCTATCTCTACACAGCAGCGTGAATGATAAAGGGCTGAGA 1476
Qy 597 ThrLysGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrPro 616
Db 1477 GTCCTGGAGAGAC---AATTGAGCTTTGGATAAAATG-----CTGCAAGATGTCAGATG 1527
Qy 617 ProGlu-----IlePheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMet 633
Db 1528 CCTCCAAAAAAGCTGAGCTCATTTTGAATGAACAAATAATTT---TGGTATCAGATG 1584
Qy 634 LeuTyrLysProHisAspLeuGlnProGlyLysTyrProThrValLeuPheIleTyr 653
Db 1585 ATCTGGCTCCTCAT---TTTGATAAATCCAGAAATATCCTCTACTATTAGATGTGAT 1641

Qy 654 GlyGlyProGlnValGlnLeuValAsnAsnArgPheLysGlyValLysTyrPheArgLeu 673
Db 1642 GCAGGGCCCATGTAGTCAA-----AAAGCAGACACTGTCTTTCAGACTG 1683
Qy 674 Asn-----ThrLeuAlaSerLeuGlyTyrValValValVal---IleAspAsnArg 689
Db 1684 AACTGGGCCACTTACCTTGCAGACAGACAGAAACATTATAGTAGCTAGCTTTGATGGCAGA 1743
Qy 690 GlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLysTyrLysMetGlyGlnIle 709
Db 1744 GGAAGTGTATTACACAGCAGATAAGATCATGCAATCAACAGAAAGACTGGGACAACTTT 1803
Qy 710 GluIleAspAspGlnValGluGlyLeuGlnTyrLeuAlaSerArgTyrAspPheIleAsp 729
Db 1804 GAAGTTGAAGATCAAAATTTGAAGCAGCAGCAAAATTT---CAAAAATGGGATTTGTGAC 1860
Qy 730 LeuAspArgValGlyIleHisGlyTyrTrpSerTyrGlyTyrLeuSerLeuMetAlaLeu 749
Db 1861 AACAAAGCAATTCGAATTTGGGGCTGGTCATATGAGGGTACGTAACCTCAATGCTCTG 1920
Qy 750 MetGlnArgSerAspIlePheArgValAlaIleAlaGlyAlaProValThrLeuTrpIle 769
Db 1921 GGATCGGAAAGTGGCTGTTCAAGTGTGAATAGCGGTGGCGCTGTATCCCGGTGGAG 1980
Qy 770 PheTyrAspThrGlyTyrThrGluArgTyrMetGly-----HisProAspGlnAsnGlu 787
Db 1981 TACTATGACTGCTGATACACAGACGTTACATGGGTCTCCCAACTCCACAGACAACTT 2040
Qy 788 GlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGluLysPheProSerGluProAsn 807
Db 2041 GACCAATTACAGAAATTCACAGCTGATGACAGAGCTGAAATTTT-----AACAGATT 2094
Qy 808 ArgLeuLeuLeuLeuHisGlyPheLeuAspGluAsnValHisPheAlaHisThrSerIle 827
Db 2095 GAGTACCTCTCTATTATTCATGGAACAGCAGATGATACGTTCACTTTTCAGCAGCTCAG 2154
Qy 828 LeuLeuSerPheLeuValArgAlaGlyLysProTyrAspLeuGlnIleTyrProGlnGlu 847
Db 2155 ATCTCCAAAGCCCTGGTGGATGTTGGATGGGATTTCCAGGCAATGTGTATATCTATG 2214
Qy 848 ArgHisSerIleArgValProGluSerGlyGluHisTyrGluLeuHisLeuLeuHisTyr 867
Db 2215 GACCATGGAATAGTAGTACGACAGCAGCAACCAACATATATATATACCATGACCTTC 2274
Qy 868 LeuGlnGlu 870
Db 2275 ATAAACAA 2283

RESULT 7

US-10-501-035-34
Sequence 34, Application US/10501035
Publication No. US20060046249A1

GENERAL INFORMATION:

APPLICANT: Bristol-Myers Squibb Company.

TITLE OF INVENTION: IDENTIFICATION OF POLYNUCLEOTIDES AND POLYPEPTIDE FOR PREDICTING
TITLE OF INVENTION: ACTIVITY OF COMPOUNDS THAT INTERACT WITH PROTEIN TYROSINE KINASES
TITLE OF INVENTION: AND/OR PROTEIN TYROSINE KINASE PATHWAYS

FILE REFERENCE: D0185 PCT

CURRENT APPLICATION NUMBER: US/10/501,035

CURRENT FILING DATE: 2004-07-09

PRIOR APPLICATION NUMBER: US 60/350,061

PRIOR FILING DATE: 2002-01-18

NUMBER OF SEQ ID NOS: 795

SOFTWARE: PatentIn version 3.2

SEQ ID NO 34

LENGTH: 3407

TYPE: DNA

ORGANISM: Homo sapiens

US-10-501-035-34

Alignment Scores:

Pred. No.: 5.4e-46

Score: 529.00

Length:

3407

Matches:

186

Percent Similarity: 42.0%	Conservative: 109				
Best Local Similarity: 26.5%	Mismatches: 280				
Query Match: 11.3%	Indels: 128				
DB: 9	Gaps: 32				
US-10-825-632-1 (1-882) x US-10-501-035-34 (1-3407)					
QY	206	ProAlaAspProAspTrrpille-----AlaPheIleHis 216	Db	1288	ATCGGG-----ATAGAAGCTCTAACCAAGTGATTATCTATCTACTACATTAGT 1332
Db	520	CCAAACAACACACAGTGGGTCCACATGGTCCACAGTGGGTGCATATAAATGGCATATGTTGG 579	QY	526	GlyThr---LysAspSerProLeuGluHisHisLeuTyrValValSerTyrValAsnPro 544
QY	217	SerAsnAspIleTrrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrVal 236	Db	1333	AATGAATATAAAGGAATCCAGGAGGAAGGAATCTTTATAAAATCCAACTTATTGTACTAT 1392
Db	580	AACATGACATTTATGTTAAATTTGAACCAAAATTTACCAAGTTTACAGAAATCACATGGACG 639	QY	545	GlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCysIleSer---Gln 563
QY	237	HisAsnGluLeuAlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheVal 256	Db	1393	ACAAAGTGACATGCCTC-----AGTTGTGAGCTGAATCCGGAA 1431
Db	640	GGG-----AAAGAAGATATAATATAATGAATAAATACTGACTGGGTT 681	QY	564	HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHis-----Cys 580
QY	257	LeuGlnGluGlu---PheAspArgTyrSerGlyTyrTrrpCysProLysAlaGluThr 275	Db	1432	AGGTGTCAGTACTATCTCTGTCATTCAGTAAAGGCGGAGATATTATCAGCTGAGATGT 1491
Db	682	TATGAAGAGGAGTCTTCAGTGCCTACTCTGCTGCTGCTGCTGCTCCAAACGGCACTTTT 741	QY	581	-----ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLys 596
QY	276	ThrProSerGlyGlyLysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluVal 295	Db	1492	TCCGCTCCTGCTCGCCCTCTATCTCTACACAGCAGCGTGAATGATATAAGGCGTGAGA 1551
Db	742	-----TTAGCATATGCCCAATTTAAGCAGACAGATC 774	QY	597	ThrLysGluPheTrrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrPro 616
QY	296	GluIleIleHis-----ValThrSerProMetLeuGluThrArgArgAlaAspSer 312	Db	1552	GTCTCTGGAAGAC---AATTCAGCTTTGGATAAAATG-----CTGCAGAAATGTCCAGATG 1602
Db	775	CCACTTATTGAATACTCTCTTACTCTGATGAGTCACTGCAGTACCCAAAGACTGTACGG 834	QY	617	ProGlu-----IlePheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMet 633
QY	313	PheArgTyrProLysThrGlyThrAlaAsnProLysValThrPheLysMetSerGluIle 332	Db	1603	CCCTCCAAAAAACTGGACTTCATTATTTGAATGAACAAAAATTT---TCGTATCAGATG 1659
Db	835	GTTCATATCCAAAGCAGGAGCTGTGAATCCAACTGTAAAGTTCTTTGTT-----GTA 888	QY	634	LeuTyrLysProHisAspLeuGlnProGlyLysLysTyrProThrValLeuPheIleTyr 653
QY	333	MetIleAspAlaGluGlyArgIleIleAspValIleAspLysGluLeuIleGlnProphe 352	Db	1660	ATCTTGCTCTCTCAT---TTTGATAAAATCCAAGAAATATCTCTACTATTAGATGTGTAT 1716
Db	889	AATACAGACTCTCTCAGCTCAGTCACCAATCCAACTTCCATACAAATCACTGCTCTGCT 948	QY	654	GlyGlyProGlnValGlnLeuValAsnAsnArgPheLysGlyValLysTyrPheArgLeu 673
QY	353	GluIleLeuPheGluGlyValGluTyrIleAlaArgAlaGlyTrrpThrProGluGlyLys 372	Db	1717	GCAGGCCCATGTAGTCAA-----AAAGCAGACACTGCTTTCAGACTG 1758
Db	949	TCTATGTTGATA---GGGGATCACTACTTGTGTGTGTGATGCATGGCA-----993	QY	674	Asn-----ThrLeuAlaSerLeuGlyTyrValValValVal---IleAspAsnArg 689
QY	373	TyrAlaTrrpSerIleLeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSer 392	Db	1759	AACTGGGCCACTTACTCTTGCAGCACAGAAACATTTATAGTAGTACTGTTCATGCGAGA 1818
Db	994	-----ACACAAGAAAGAAATTTCTTTGCACTGGCTC---1023	QY	690	GlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLysTyrLysMetGlyGlnIle 709
QY	393	ProGluLeuPheIleProValGluAspValMetGluArgGlnArgLeuIleGluSer 412	Db	1819	GGAGTGGTTACCAAGGAGATAGATCATGTCATCAACAGAGAACTGGGAACATTT 1878
Db	1024	-----AGGAGGATTCAGAAC 1038	QY	710	GluIleAspAspGlnValGluGlyLeuGlnTyrLeuAlaSerArgTyrAspPheIleAsp 729
QY	413	ValProAspSerValThrProLeuIleIleTyrGluGluThrThrAspIleTrrpIleAsn 432	Db	1879	GAAGTTGAAGATCAATTTGAAGCCAGCCAGACAATTT---TCAAAAATGGGATTTGTGGAC 1935
Db	1039	TAT-----TCGGTCATGGATATTTGTGACTATGATGATCCAGTCCAGAGATGG---AAC 1089	QY	730	LeuAspArgValGlyIleHisGlyTrrpSerTyrGlyGlyTyrLeuSerLeuMetAlaLeu 749
QY	433	IleHisAspIlePheHisValPheProGlnSerHisGluGluGluIleGluPheIlePhe 452	Db	1936	AACAAACGAATTTGCGGCTGGTCATATGAGGGTACGTAAACCTCAATGCTCTG 1995
Db	1090	TGC-----TTAGTGGCAGGCAACACATTTGAATGACTACTGCTGGCTGG 1134	QY	750	MetGlnArgSerAspIlePheArgValAlaIleAlaGlyAlaProValThrLeuTrrpIle 769
QY	453	AlaSerGluCysLysThrGlyPheArgHis-----LeuTyrLys 465	Db	1996	GGATCGGGAAGTGGCGTGTTCAGGTGTGGAATAGCGGTGGCGCTGTATCCCGTGGGAG 2055
Db	1135	GTTCGAAGATTTAGGCTTCAGAACCTCATTTTACCTTGATGTAATAGCTTCTACAG 1194	QY	770	PheTyrAspThrGlyTyrThrGluArgTyrMetGly-----HisProAspGlnAsnGlu 787
QY	466	IleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGlyGlyLeuProAlaPro 485	Db	2056	TACTATGACTCAGTGTACACAGAACGTTACATGGGTCTCCCACTCCAGAGCAACACCTT 2115
Db	1195	ATCATCAGC-----AATGAAGAAGGTTACAGACACTT-----1227	QY	788	GlnGlyTyrTrrpLeuGlySerValAlaMetGlnAlaGluLysPheProSerGluProAsn 807
QY	486	SerAspPheLysCysProIleLysGluGluIleAlaIleThrSerGlyGluTrrpGluVal 505	Db	2116	GACCATTTACAGAAATTTCAACAGTCATGACAGCAGCTGAAAAATTTT---AAACAGTT 2169
Db	1228	TGCTATTTCCAAATAGATAAAAGACTGCACATTTATTACAAAGGCACTGGGAAGTC 1287	QY	808	ArgLeuLeuLeuLeuHisGlyPheLeuAspGluAsnValHisPheAlaHisThrSerIle 827
QY	506	LeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArgLeuValTrrpPheGlu 525	Db	2170	GAGTACTCTCTTATTCATGGAACAGCAGATGATAACGTTCTACTTTTCAGCAGCTCAGCTCAG 2229
Db			QY	828	LeuLeuSerPheLeuValArgAlaGlyLysProTyrAspLeuGlnIleTrrpProGlnGlu 847
			Db	2230	ATCTCCAAAGCCCTGGTTCGATGTTGGAGTGGATTTCCAGGCAATGTGGTGTATCTGATGAA 2289
			QY	848	ArgHisSerIleArgValProGluSerGlyGluHisTyrGluLeuHisLeuLeuHisTyr 867
			Db	2290	GACCATGGAATAGTAGCAGCAGCACACCAACATATATATACCCACATGAGCCACTTC 2349

Qy 868 LeuGlnGlu 870
: : : : : : :
Db 2350 ATAAACAA 2358

RESULT 8

```

US-10-522-789-1
; Sequence 1, Application US/10522789
; Publication No. US20050260732A1
; GENERAL INFORMATION:
; APPLICANT: TANABE SEIYAKU CO., LTD.
; TITLE OF INVENTION: Three-dimensional structure of dipeptidyl peptidase IV
; FILE REFERENCE: 03-039-PCT
; CURRENT APPLICATION NUMBER: US/10/522,789
; CURRENT FILING DATE: 2005-01-28
; PRIOR APPLICATION NUMBER: US 60/398,761
; PRIOR FILING DATE: 2002-07-29
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 2301
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2301)
; OTHER INFORMATION:
US-10-522-789-1

```

Alignment Scores:		
Pred. No.:	8,486-46	Length: 2301
Score:	\$25.00	Matches: 186
Percent Similarity:	41.6%	Conservative: 108
Best Local Similarity:	26.3%	Mismatches: 277
Query Match:	11.2%	Indels: 136
DB:	8	Gaps: 33

US-10-825-632-1 (1-882) x US-10-522-789-1 (1-2301)

206	ProAlaAspProAspTrpIle-----AlaPheIleHis	216
	::: :::	
445	CCAAACACACACACAGCTGGGTCCACATGGTCCACAGTGGGTCCATAAAATGGCATATTTGGG	504
217	SerAsnAspIleTrpIleSerAsnIleValThrArgGluGluArgGluLeuThrTyVal	236
::: :::	::: :::	
505	AACAATGACATTATGTATAAATGAACCAAAATTTACAAAGTTACAGAAATCACATGGAGG	564
237	HisAsnGluLeuAlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheVal	256
::: :::	::: :::	
565	GGG-----AAAGAGATATAATATATATATGAATAACTGACTGGGTT	606
257	LeuGlnGluGlu---PheAspArgTyrSerGlyTyrTrpTrpCysProIysAlaGluThr	275
::: :::	::: :::	
607	TATGAGAGGAAAGCTCTTCAGTGCCTACTCTGCTGTGGTGGTCTCCAAACGCGCACTTTT	666
276	ThrProSerGlyGlyValIleLeuArgIleLeuTyrgLluAsnAspGluSerGluVal	295
::: :::	::: :::	
667	-----TTAGCATATGCCCAATTTAACGACACAGAAATC	699
296	GluIleIleHis-----ValThrSerProMetLeuGluThrArgArgAlaAspSer	312
::: :::	::: :::	
700	CCACTTATTTGAATACTCCTTCTCTGATGAGTCACTGCAGTACCCAAAGACTGTACGG	759
313	PheArgTyrProIythrGlyThrAlaAsnProIysValThrPheIysMetSerGluIle	332
::: :::	::: :::	
760	GTTCATATATCCAAAGCGAGGAGCTGTGAATCCCACTGTAAAGTCTTTGTT-----GTA	813
333	MetIleAspAlaGluGlyArgIleIleAspValIleAspLysGluLeuIleGlnProPhe	352
::: :::	::: :::	
814	AATACAGACTCTCTCAGCTCAGTCACCAATGCAACTTCCATACAAATCACTGCTCTGCT	873
353	GluIleLeuPheGluGlyValGluTyrIleAlaArgAlaGlyTrpThrProGluGlyLys	372
::: :::	::: :::	

Db 1135 TGCACCTC----- 1143
Qy 484 AlaProSerAspPheLys---CysProIleLysGluGluLeuAlaIleThrSerGlyGlu 502
Db 1144 ---CCGAAGATAAGAAAGACTGTACATTT-----ATTACAAAGAGGCC 1185
Qy 503 TrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArgLeuVal 522
Db 1186 TGGGAAGTCAAT-----AGTATCGAAGCTCTGACCGAGATTATCTATAC 1230
Qy 523 TyrPheGluGlyThr---LysAspSerProLeuGluHisGlySerValValSerTyr 541
Db 1231 TACATTAGTAAACCAATATAAGAAATGCCAGAGAAAGAAATCTCTATAAAAT----- 1284
Qy 542 ValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCysIle 561
Db 1285 -----CAACTTACTGAC-----CACAAATGTGAAGTGCCTT 1317
Qy 562 Ser-----GlnHisCysAspPhePheLeuSerLysTyrSerAsnGlnLys 576
Db 1318 AGTTGTGACCTGAATCCAGAAAGATGTCAGTATTATGCGGTATCATTTAGTAAAGAGGCA 1377
Qy 577 AsnProHis-----Cys-----ValSerLeuTyrLysLeuSerSerPro 589
Db 1378 AAGTACTATCAGCTGGGATGTTGGGGCCCGGCTGCGCCCTCTACACTCTACATCTAGC 1437
Qy 590 GluAspAsp-ProThrCysLysThrLysGluPheTrpAlaThrIleLeuAspSerAlaG 609
Db 1438 ACGGATCAATAAGAGCTGC-----GAGTCTGGAGAACATTCGTCT-TTGATAG 1487
Qy 609 yProLeuProAspTyrThrProProGlu-----IlePheSerPheGluSerThr 626
Db 1488 AATGCTGCAGGATGCCAGATGCTTCAAAAAAATGAGCTTCATCTGTTTGAATCAAAAC 1547
Qy 626 rGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeuProGlyLysLys 646
Db 1548 AAGATTTT---TGATATCAATGATCTTCCGCCCTCAT---TTTGATAAATCCAGAAATA 1601
Qy 646 rProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeuValAsnAsnArgPhe 666
Db 1602 TCCTCTACTATTAGATGATATGAGGTCCTGATGTCATCA-----AA 1643
Qy 666 sGlyValLysTyrPheArgLeuAsn-----ThrLeuAlaSerLeuGlyTyrVal 683
Db 1644 AGCAGATGCTTCTTCAGACTGAACTGGGCCCACTTACCTTGCAGATACAGAAACATCAT 1703
Qy 683 lValVal---IleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAla 702
Db 1704 AGTAGTAGCTTTGACGGCAGAGAGAGTGGTTTACCAAGGAGATAAGATCATGATGCAAT 1763
Qy 702 eLysTyrLysMetGlyGlnIleGluLeuAspAspGlnValGluGlyLeuGlnTyrLeu 722
Db 1764 CAACAGAGATTGGGAACACTGGAAGTTGAGATCAATGATGACACCCAGGCAATTTGT 1823
Qy 722 aserArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyrGly 742
Db 1824 A---AAATGGGATTTGTGGTAGCAGCAGGATTCATTTGGGCGCTGTATATGAGG 1880
Qy 742 yTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIleAla 762
Db 1881 GTATGTAACCTCAATGGTCTCGGATCGGGAAGTGGCGTTCAGTTCGGAATAGTGT 1940
Qy 762 yAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMetGly-- 781
Db 1941 GGCACCTGTGTACCGTGGGAGTACTATGACTGTGTACAGGCGTTACATGGGTCT 2000
Qy 782 ----HisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAla 800
Db 2001 CCCAATTCAGAGACACCTTGACCATTTACAGGAACTCAACAGTCTATGACGAGCTGA 2060
Qy 800 uLysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluLeuVal 820
Db 2061 ACATTTT-----AAACAAGTTGAGTACTCTCTTCTTATTCATGGAACGGCAGATGATATGT 2114

Qy 820 lHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAs 840
Db 2115 TCACCTTTCAGCAGTCAGCTCAGATCTCCAAAGCCCTGGTGTGATCTGGTGTGGATTTCCA 2174
Qy 840 pLeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHis 860
Db 2175 AGCAATGTGGTACAGGATGAAGACACGCGGATCGTAGCAGCAGCTACACGACAT 2234
Qy 860 rGluLeuHisLeuLeuHisTyrLeuGlnGlu 870
Db 2235 CTATTTCCACATGAGCCATTTCTCTCCAGCAG 2265

RESULT 10

US-10-932-182A-5649
; Sequence 5649, Application US/10932182A
; Publication No. US20060046253A1
; GENERAL INFORMATION:
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHIISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: 030685-043
; CURRENT APPLICATION NUMBER: US/10/932, 182A
; CURRENT FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 5649
; LENGTH: 2778
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-10-932-182A-5649

Alignment Scores:
Pred. No.: 6,89e-42 Length: 2778
Score: 490.00 Matches: 183
Percent Similarity: 41.3% Conservative: 124
Best Local Similarity: 24.6% Mismatches: 316
Query Match: 10.4% Indels: 120
Gaps: 30

US-10-825-632-1 (1-882) x US-10-932-182A-5649 (1-2778)

Qy 166 AlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGlyPheThrGlnGlnPro 185
Db 796 TCTTCCAAAGGCTATTATTGGATTAGGATTTGAGTACAGCT-----ACTATCGAACCG 849
Qy 186 LeuArgProAsnLeuValGluThrSerCysProAsnIleArgMetAspProLysLeuCys 205
Db 850 ATTCTACCACCG--GATGAAGCAGCTGATGGAATGTACATGGTTATTGAAACTGTCC 906
Qy 206 ProAla-----AspPro-----AspTrpIleAlaPheIleHisSerAsnAspIleTrp 221
Db 907 TATGCCCATTTCTCACCTCTTACAATTACATTTATTGTTGCTAGAGAATAATTGTTT 966
Qy 222 lLeSerAsnIleValThrArgGluGluArgLeuThrTyr---ValHisAsnGluLeu 240
Db 967 CTGCAACAGGGGAATTCAGATACGCCAAGAGCTTACCCTATGATGGCTCAAAAGACATA 1026
Qy 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu 260
Db 1027 TTTTAATGCTAAACACAT-----TCGATATACGAGCAAGAG 1062
Qy 261 ----PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGly 279
Db 1063 GTCTCGCGTCCCAACCAAGCTATATGTTGG-----GCCCCAGATGAC 1104
Qy 280 GlyValIleuArgIleLeuTyrGluLeuAsnAspGluSerGluValGluIleIleHis 299
Db 1105 TCAAAAGCTGTTTTCGAAGATTTAATGACCCCGCTTGAAGATATTCAACTAACAC 1164

QY	300	ValThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGly	319	Db	2044	AAACACGCAACCTTAGATCCCAAGAAATATCCGATATTAGTAAACATTTATGGTGGC	2103
Db	1165	TATACGACACGAGGAACCATATCTTCCAGTTCTAAATAAATATCTCTAAACACGAG	1224	QY	656	ProGlnValGlnLeuValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThr	675
QY	320	ThrAlaAsnProLysValThrPheLysMetSerGluLeuMetIleAspAlaGluGlyArg	339	Db	2104	CCAGGATCTCAATCATTCACCACT-----AAATCGCTCGCTGCATTCGAACAGTCTGT	2157
Db	1225	TCTAAACCCCAAGTTTGATTTA-----TTTTTAATCAACTTACAGAACGAG	1272	QY	676	LeuAlaSerLeuGlyTyrValValValIleAspAsnArgGlySerCysHisArgGly	695
QY	340	IleIleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluVal	359	Db	2158	GCCTTCAGGATTGGATGCAATCATCTCCAAATAGAACCAAGAGGAACCTGGGGTAAAGT	2217
Db	1273	GTACTTTTATTCAGTCATATACGGGT---GGTCAAAAGGATTCATCTTTGTATAATGGG---	1326	QY	696	LeuLysPheGluGlyAlaPheLysTyrLysMetGlyGlnIleGluIleAspGlnVal	715
QY	360	GluTyrIleAlaArgAlaGlyTyrThrProGluGlyLysTyrAlaTyrSerIleLeuLeu	379	Db	2218	TGGAGTTTCAGATCTTCGGCTCGCGGTAGATTGGGATATTGGAGCGCTAGGATATTATT	2277
Db	1327	-----AAATGGATAGTCCAGATTCATTTAGATTTCGAATTT---ACG	1365	QY	716	Glu---GlyLeuGlnTyrLeuAlaSerArgTyrAspPheIleAspLeuAspArgValGly	734
QY	380	AspArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProVal	399	Db	2278	GAGGTACCAAGAAATTTTACAAAGAACCAACACATATAGACGAAACAGGTCGCA	2337
Db	1366	GACAGAAATCCAAAGGTGTGGACGTT-----	1392	QY	735	IleHisGlyTyrSerTyrGlyTyrLeuSerLeuMetAlaLeu---MetGlnArgSer	753
QY	400	GluAspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrPro	419	Db	2338	ATTTCGGGTTCGTCTGCTGCTGCTTTGTTAGTTTGAAGACCGTAGAATTGGATTAATGTT	2397
Db	1393	-----AAAGTATACGATGTACCGCTCTTCT-----	1416	QY	754	AspIlePheArgValAlaIleAlaGlyAlaProValThrLeuTyrIlePheTyrAspThr	773
QY	420	LeuIleIleTyrGluGluThrThrAspIle-----TyrIle---AsnIle	433	Db	2398	GAAACATTCAAATACACATGGCTGTGCACCACTGAACTGGAACCTGTATGACTCC	2457
Db	1417	AAACTACTTACTGTGAGAACCATCTGATTCCGAATTCGTTCAATGGATGGATTGAAAAGACC	1476	QY	774	GlyTyrThrGluArgTyrMetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGly	793
QY	434	HisAspIlePheIleValPheProGlnSerHisGluGluGluIleGlu-----PheIle	451	Db	2458	ATCTACACGAAAGGTATATGAATCAACCATCGGAAATGAGAGGGTTATTTTGACATA	2517
Db	1477	AAAGATATCTTAAGAAATCCACCAAAACCCGAAATCCGAAGAACGAGTACCGGATATATC	1536	QY	794	SerValAlaMetGlnAlaGluLysPheProSerGluProAsnArgLeuLeuLeuHis	813
QY	452	PheAlaSerGluCysLeYThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLys	471	Db	2518	TCTACTATTAATAATTCACAGTCTTCGAATCG---CTGAAGCGATCTCTTTGTATGCAC	2574
Db	1537	GATATTCACGGCGGACATAACGGGTTCATCACTCTTTTATTCACCACTCTTT---	1593	QY	814	GlyPheLeuAspGluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeuVal	833
QY	472	GluSerLysTyrLysArgSerSerGlyLeuProAlaProSerAspPheLysCysPro	491	Db	2575	GGGACCTCTGATGATAATGTTTCATATTCAAAATACGTTTAGATTGTCGATCAACTGAAT	2634
Db	1593	-----	1593	QY	834	ArgAlaGly---LysProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleArg	852
QY	492	IleLysGluGluIleAlaIleThrSerGlyGluTyrGluValLeuGlyVargHis-----	509	Db	2635	CTTTTGGGATTGGCCAAATATGACATCATATTTTCCCGATTCAGATCATTTCTATCAGA	2694
Db	1594	GCCAAAGATGGCATTCATGCTTAAGGAAACTGGGAGTTACTGGGAACGGTATATTT	1653	QY	853	ValProGluSerGlyGluHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeu	872
QY	510	GlySerAsnIleGlnValAspGluValArgLeuValTyrPheGluGlyThrLysAsp	529	Db	2695	CACCACACGCTCAAAAGATTGTATTCAGAAACTGTACTACTGTTAAAGGATGCCTTT	2754
Db	1654	GGTTAGCAATATGAACCGGAT-----ACCGTATTTTTCACGCCAATGAGATT	1701	QY	873	GlySerArg	875
QY	530	SerProLeuGluHisIleLeuTyrValValSerTyrValAsnProGlyGluValThrArg	549	Db	2755	GACGGCCGC	2763
Db	1702	GGAAACCATGTCACACACCTTTATAGTACCAGTTT---AATCTGCACTATGAGCAAA	1758	RESULT 11			
QY	550	Leu-----ThrAspArgGlyTyrSerHisSerCysCysIleSerGln	563	US-10-932-182A-5649			
Db	1759	CTTCAAGTTATGCAAAATCCATCTGATAAATATATATATTTTATGACTTTTGAATTTAGCTCT	1818	/ Sequence 5649, Application US/10932182A			
QY	564	HisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCysValSer---	582	/ Publication No. US20060046253A1			
Db	1819	AGTGAAGATATGCCATTTCAAAGAAACTAGGCCCGGTGACCACTGGAAGTAGTGGGA	1878	/ GENERAL INFORMATION YOSHIHIRO			
QY	583	-----LeuTyrLysLeuSerSerProGluAspAspProThr-----Cys	595	/ APPLICANT: NAKAO, YOSHIHIRO			
Db	1879	CCGCTTACGGAAGTACTGATCTTACCAATGCTAAACGGCATTCGGTGTGTCAGCTGACC	1938	/ APPLICANT: NAKAMURA, NORIHIISA			
QY	596	LysThrLysGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThr	615	/ APPLICANT: KODAMA, YUKIKO			
Db	1939	AGGATCGAAATTTTAAGAGACAAATAAGGGTTACGATCTACCAATTCACGATCAAG	1998	/ APPLICANT: FUKIMURA, TOMOKO			
QY	616	ProProGluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyr	635	/ APPLICANT: ASHIKARI, TOSHIHIKO			
Db	1999	-----TCAATGGTTCTAAGTGATGGTATGGAATAAATACATACGAAATT	2043	/ TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS			
QY	636	LysProHisAspLeuGlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGly	655	/ FILE REFERENCE: 030685-043			
				/ CURRENT APPLICATION NUMBER: US/10/932,182A			
				/ CURRENT FILING DATE: 2004-09-02			
				/ NUMBER OF SEQ ID NOS: 197023			
				/ SOFTWARE: PatentIn version 3.3			
				/ SEQ ID NO 5649			
				/ LENGTH: 2778			
				/ TYPE: DNA			
				/ ORGANISM: Saccharomyces pastorianus			
				US-10-932-182A-5649			

Qy	814	GlyPheLeuAaspGluAenValHisPheAlaHisThrSerIleLeuLeuSerPheIleuVal	833	
Db	2575	GGGACCTCTGATGATAATGTTCAAAATACGTTTAGATTGTGATCACTGAAT	2634	
Qy	834	ArgAlaGly---LysProTyrAaspLeuGlnIleTyrProGlnGluArgHisSerIleArg	852	
Db	2635	CTTTTGGGATTGGGCAATTATGACATGATATTTTCCCGATTACAGATCATCTATCAGA	2694	
Qy	853	ValProGluSerGlyGluHisTyrGluLeuHisLeuLeuHisTyrIleuGlnGluAenLeu	872	
Db	2695	CACCACAACGCTCAAAAGATTGTATTCCAGAAACTGTACTACTGGTTAAAGGATCCCTTT	2754	
Qy	873	GlySerArg	875	
Db	2755	GACGGCCGC	2763	
RESULT 12				
US-11-136-527-2130				
; Sequence 2130, Application US/11136527				
; Publication No. US20050287570A1				
; GENERAL INFORMATION:				
; APPLICANT: Wyeth				
; TITLE OF INVENTION: Mounts, William M				
; FILE REFERENCE: 031896-041000 (AM101086)				
; CURRENT APPLICATION NUMBER: US/11/136,527				
; PRIOR FILING DATE: 2005-05-25				
; PRIOR APPLICATION NUMBER: US 60/574,294				
; NUMBER OF SEQ ID NOS: 362830				
; SOFTWARE: Patent in version 3.2				
; SEQ ID NO 2130				
; LENGTH: 4852				
; TYPE: DNA				
; ORGANISM: Rattus norvegicus				
US-11-136-527-2130				
Alignment Scores:				
Pred. No.:	2,65e-40	Length:	4852	
Score:	478.50	Matches:	183	
Percent Similarity:	40.5%	Conservative:	101	
Best Local Similarity:	26.1%	Mismatches:	297	
Query Match:	10.2%	Indels:	121	
DB:	14	Gaps:	30	
US-10-825-632-1 (1-882) x US-11-136-527-2130 (1-4852)				
Qy	206	ProAlaAaspProAaspTrrIle	216	-----AlaPheIleHis
Db	527	CCAAATAATACACAGTGGATCACATGGTCAACAAGGTCACAMATTGGCATATGTCGG	586	
Qy	217	SerAsnAaspIleTrrIleSerAsnIleValThrArgGluGluArgArgLeuThrVal	236	
Db	587	AAGAAATGATTTATGTTTAAATTCGAACACACATTTGCTAGTCAGAGAYCACA	640	
Qy	237	HisAsnGluLeuAlaAsnMetGluGluAaspAlaArgSerAlaGlyValAlaIleThrPheVal	256	
Db	641	-----TCACAKGAAGAAGAAATGTWATATTAAACGGWATAAATGACTGGTT	688	
Qy	257	LeuGlnGluGlu---PheAaspArgTyrSerGlyTyrTrrTrrCysPProLysAlaGluThr	275	
Db	689	TATGAAGAGAGAATCTTCGGTCCCTACTCCGACCTGGTG	730	
Qy	276	ThrProSerGlyGlyLysIleLeuArgIleLeuTyrGluGluAenAaspGluSerGluVal	295	
Db	731	TCCTCCAAAGGCACTTTCTAGCTTATGCCAGTTTAAACGACACCGGAGTCCCTCTCAT	790	
Qy	296	GluIleIleHisValThrSerProMetLeuGluThrArgArgAlaAaspSerPheArgTyr	315	
Db	791	GAATACTCTCTACTCTGATGAGTCACTGCGAGTACCCCAAGACAGTSTGGATTCGGTAC	850	
Qy	316	ProLysThrGlyThrAlaAsnProLysValThrPheLysMetSerGluIleMetIleAasp	335	


```
Db 1349 -AAAGACACTGTGGAAATGCTATTCAAATTACAGTGGCAAGTGGAGGCCATA----- 1402
Qy 508 gHisGlySerAsnIleGlnValaspGluValargLeuValTyrPheGluGlyThrLy 528
Db 1403 -----AATATA-----TTACAGATAACACA 1422
Qy 528 sAspSerProLeuGluHisHisLeuTyrValValSerTyrValAsnProGlyGluValTh 548
Db 1423 GGATTTCATGTTTATTCTAGCAATGAATTTGAAGAAATACCTCGGAAGAGAAACATCTA 1482
Qy 548 rArgLeuThrAspArgGlyTyrSerHisSer---CysCysIleSerGlnHis----- 564
Db 1483 CAGAATTAGCAATTGGAAGCTATCTCTCAAGCAAGAAGTGTGTACTTCTGCCATCTAGGAA 1542
Qy 565 -----CysAspPhePheIleSerTyrSerAsnGlnLysAsnProHisCysValSe 582
Db 1543 AGAAAGGTGCCAATATTACACAGCAAGTTTCAGCGACTACGCCAAGTACTATGCATTTGT 1602
Qy 582 rLeuTyr-----LysLeuSerSerProGluAspAspProThrCysLysThrLy 598
Db 1603 CTGCTACGGCCAGGCATCCCATTTCCACCTTCATGATGGAGCGACTGATCAAGAA-- 1660
Qy 598 sGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAsp----- 613
Db 1661 -----ATTAAATCTCGNAGNAAACAGGAATTGGAAATGCTTTGAAAATAT 1710
Qy 614 -TyrThrProGluLePheSerPheGluSerThrThrGlyPheThrLeuTyrGlyWe 633
Db 1711 CCAGCTGCCTAAAGAGGAATTAAGAACTTGGAAGTAGATGAATTAATCTTTATGTACAA 1770
Qy 633 tLeuTyrLysProHisAspLeuGlnProGlyLysLysTyrProThrValLeuPheIleTy 653
Db 1771 GATGATCTCTCTCTCTCAATTGACAGATCAAGAAAGTATCCCTGCTAATTCAGATGTA 1830
Qy 653 rGlyGlyProGlnValGlnLeuValAsnAsnArgPheLysGlyValLysTyrPheArgLe 673
Db 1831 TGGTGGTCCCTCGAGTCAGAGTGAAGTCT-----GTATTTTGGCTGT 1872
Qy 673 uAsnThrLeuAlaSerLeu-----GlyTyrValValValIleAspAsnAr 689
Db 1873 TAATTTGGATATCTTATCTTGCAGTAAGGAAGGATGTCATTCCTGCTGGTGGTGGT 1932
Qy 689 gGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLysTyrLysMetGlyGlnI 709
Db 1933 AGGAACAGCTTCCAAAGTGCAGAACTCCTCATGCAGTGTATCGAAAGCTGGTGTTA 1992
Qy 709 eGluLeAspAspGlnValGlyLeuGlnTyrLeuAlaSerArgTyrAspPheIleAs 729
Db 1993 TGAAGTTGAGACCCAGATTACAGCTGCAGAAATTCATA--GAAATGGGTTTCATTGA 2049
Qy 729 pLeuAspArgValGlyIleHisGlyTyrSerTyrGlyGlyTyrLeuSerLeuMetAla 749
Db 2050 TGAAGAAAGATAGCCATATGGGCTGTCTATGGAGATACGTTTTCATCTGCTGCTCT 2109
Qy 749 uMetGlnArgSerAspIlePheArgValAlaIleAlaGlyAlaProValThrLeuTrpI 769
Db 2110 TGCATCTGGAAGTGTCTTTCAAATGTGGTATAGCAGTGGCTCCAGTCTCCAGCTGGGA 2169
Qy 769 ePheTyrAspThrGlyTyrThrGluArgTyrMetGlyHisPro-----AspGlnAsnGl 787
Db 2170 ATATTACGGCTCTGTCTACACAGAGAGATTATGGGTCCTCCCAACAAAGAGATGATA 2229
Qy 787 uGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlyLysPheProSerGluProAs 807
Db 2230 TGAGCACTATAAGAATTCACCTGTGATGGCAAGAGCAGAAATATTTCAGAAATGTAGAC-- 2287
Qy 807 nArgLeuLeuLeuHisGlyPheLeuAspGluAsnValHisPheAlaHisThrSerI 827
Db 2288 ----TATCTTCTCATCCAGGACAGCAGATGATAATGTGCATTTTCAAAACTCAGCACA 2343
Qy 827 eLeuLeuSerPheLeuValArgAlaGlyLysProTyrAspGlnIleTyrProGlnI 847
Db 2344 GATTGCTAAAGCTCTGGTTAATGCACAGTGGATTTCAGGCAATGTGGTACTCTGACCA 2403

Qy 847 uArgHisSerIleArgValProGluSerGly-----GluHisTyrGluLeuHisLe 864
Db 2404 GAACACAGGCTTA-----TCCGGCTGTCCACGAACCACTTATACACCCACAT 2451
Qy 864 uLeuHisTyrLeuGlnGlu 870
Db 2452 GACCCCACTTCTTAAAGCAG 2470

RESULT 14
US-11-186-284-54
; Sequence 54, Application US/11186284
; Publication No. US20050268493A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Berger, Allison
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Thibodeau, Stephen N.
; APPLICANT: Burgart, Lawrence J.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF COLON CANCER
; FILE REFERENCE: MPM01-029P2RNM
; CURRENT APPLICATION NUMBER: US/11/186,284
; CURRENT FILING DATE: 2005-07-21
; PRIOR APPLICATION NUMBER: US/10/301,822
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 60/339,971
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/361,978
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/381,988
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 54
; LENGTH: 2814
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (209)...(2491)
US-11-186-284-54

Alignment Scores:
Pred. No.: 5,63e-38 Length: 2814
Score: 454.00 Matches: 177
Percent Similarity: 41.9% Conservative: 111
Best Local Similarity: 25.8% Mismatches: 278
Query Match: 9.7% Indels: 122
DB: 14 Gaps: 27

US-10-825-632-1 (1-882) x US-11-186-284-54 (1-2814)
Qy 212 IleAlaPheIleHisSerAsnAspIleTrpIleSerAsnIleValThrArgGluGluArg 231
Db 692 TTAGCATATGTCTATCAAAACAAATATCTATTGAAACAAACACCCAGGAGATCCACCTTTT 751
Qy 232 ArgLeuThrTyrValHisAsnGluLeuAlaMetGluGluAspAlaArgSerAlaGly 251
Db 752 CAAATACCATTT-----AATGGAAGAGAAATAATAATTTAATGGA 793
Qy 252 ValAlaThrPheValLeuGlnGluGluPhe-----AspArgTyrSerGlyTyrTrp 269
Db 794 ATCCCAAGACTGGGTTTATGAAGAGGAAATGCTTCTCAAAATATGCT---CTCTGGTGG 850
Qy 270 CysProLysAlaGluThrThrProSerGlyGlyLysIleLeuArgIleLeuTyrGluGlu 289
Db 851 TCTCTCT-----AATGGAATAATTTTTC-----GCATATGCGGAA 883
```


Alignment Scores:		3,69e-36	Length:	2457
Pred. No.:		436,50	Matches:	169
Score:		37.3%	Conservative:	105
Best Local Similarity:		23.0%	Mismatches:	226
Query Match:		9.3%	Indels:	235
DB:		9	Gaps:	27
US-10-825-632-1 (1-882) x US-10-932-182A-1107 (1-2457)				
QY	212	IleAlaPheIleHisSerAnAspIleTrpIleSerAsnIleValThrArgGluGluArg	231	
DB	601	ATTGCATACGTCGAAGATAACAATATATACGTCCTATCTCACAAGATCGAGAAACTACA	660	
QY	232	ArgLeuThr	245	
DB	661	CGGCTGTAACTGATGACGGAATCCCTCTCTCAAC	699	
QY	246	AspAlaArgSerAlaGlyValAlaThrPheValLeuGluGluGlu---	264	PheAspArgTyr
DB	700	GGTAAGCCGATGGTTTACGAGAGAGAGTCTTTGAGGACGAC	744	
QY	265	SerGlyTyrTrpTyrCysProLysAlaGluThrThrProSerGlyGlyLysIleLeuArg	284	
DB	745	AAGCGTGTGGTGGTCACCACTGGTGAT---	777	TAT
QY	285	IleLeuTyrGluGluAsnAspGluSerGluVal---	300	HisVal
DB	778	CTGGCGCTTTTAAAGATCGATGAATCAGAAGTTGGTGAGTTCATATTCCCTATTATGTC	837	
QY	301	ThrSerProMetLeuGluThrThrArgAlaAspSerPheArgTyrProLysThrGlyThr	320	
DB	838	CAAAATGATAAAGATGTATATCCAGAAATCGTAGCATTAAGTACCACAAAGAGTGGTACA	897	
QY	321	AlaAsnProLysValThrPheLysMetSerGluIleMetile	334	
DB	898	CCGAATCCCAT-----GCCGAGTTGGTCTACTGTATGTAAGACGAA	942	
QY	335	-----AspAlaGluGlyArgIleIleAspValIleAsp	345	
DB	943	TCATCGTACCACCTAGAGTAAATGGAGACGAAAGGTTCAAGTATATTGATTACTGAA	1002	
QY	346	LysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGluTyrIleAlaArgAla	365	
DB	1003	CTTACATGGTGGGAGATGAGGCGTTTGTAGTTAAG-----	1038	
QY	366	GlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAspArgSerGlnThrArg	385	
DB	1039	-----ACTACTGATCGTTCCTCAGATATG	1062	
QY	386	LeuGlnIleValLeuIleSerProGlu-----	394	
DB	1063	CTAACGGTATTTTGTAGTGACACGAATCCAAATCTTCATACGCTGTGAGAAATGAAAGT	1122	
QY	395	-----LeuPheIleProValGluAspAsp	402	
DB	1123	GCCATGGGGGATGGTGGAGATTACCATATACTCTCTTTATTCAGCAATGAACG	1182	
QY	403	Val---MetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeuIle	421	
DB	1183	CTTTGTAGACCTCATAATGGTTATGTCGATATCTTACCATAAATGTTTAAATCTACTTA	1242	
QY	422	IleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhePro	441	
DB	1243	GCCTACTTTGAGAATAGTAAT-----AACACACAC	1272	
QY	442	GlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPheArg	461	
DB	1272	-----	1272	
QY	462	HisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGlyGly	481	
DB	1273	-----TACAAAGTA-----	1281	

QY	482	LeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSerGly	501	
DB	1282	-----CTGCAACTGGG	1293	
QY	502	GluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArgLeu	521	
DB	1294	GATTGGGAGGTAGTT-----AATGGACCACTTGCCTTTGATTCAAAATGCGT	1344	
QY	522	ValTyrPheGluGlyThrLysAspSerProLeuGluHisIleLeuTyrValValSerTyr	541	
DB	1345	CTTTATTATTCTTACCAAAAGCACTCAACGAGCGCACTTCACTATATAGATTTC	1404	
QY	542	ValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCysIle	561	
DB	1405	CAGTCACCAACCAAGATGCATCAAGTTACTGAT-----ATT	1440	
QY	562	SerGlnHisCysAspPheIleSerLysTyrSerAsnGlnLys--AsnProHisCysVa	581	
DB	1441	TTCAAGAAGGTGTATTTCGATATATCTTTCTTCTAGGTAGGAGGTTCGGTTTAC---	1495	
QY	581	IserLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGlu---	599	
DB	1496	-----TCACCTATAAAGGACCGACGTTCCATACCACAAAATTCGTT	1536	
QY	599	-----	599	
DB	1537	GATTTTCATTACACAAAAACAGACATATATAGAAGGAAACATATTAGCGCAAACTTA	1596	
QY	600	-----PheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProProG	618	
DB	1597	TACTATTGGAGAGAAACGAAAGAAATCAGCAAA--AATTTAGAAAGATTATTTCGATTCCCAA	1655	
QY	618	uilePheSerPheGluSerThrThrGlyPheThrLeuTyrGly-----MetLeuTyrIy	636	
DB	1656	AAAAATATTCAGAGGTAAATTTGGTGAAGGATGAATCGGAGCGACATATTGGTGA	1715	
QY	636	sProHisAspLeuGlnProGly-----LysLysTyrProThrVa	649	
DB	1716	TTCTGTCGAGATCTCCCAACAGATTGTGACCAAGTTTATGATTAATGATTATTCACGTTT	1775	
QY	649	IleuPheIleTyrGlyProGlnValGlnLeuValAsnAsnArgPheLysGlyVally	669	
DB	1776	TTCTTTGATGATGTGGACCAAAATCTCAGCAATC-----ATCAA	1817	
QY	669	sTyrPheArgLeuAsn-----ThrLeuAlaSerLeuGlyTyrValValValVa	685	
DB	1818	GACATTTCTGTAGTTTCAATGAGGTCGTAGCTTCCAAATTAATGCAATTTAGTTGT	1877	
QY	685	IleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGlyAlaPheLysTyrLy	705	
DB	1878	CGTCGATGTCGTGTACTGGTTTCAAAGCCCAAGATTAGATCGGTTCTCGGTATAA	1937	
QY	705	sMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyrLeuAlaSerArgTy	725	
DB	1938	GCTTGGCGATTACGAGGCCCGTGACCAATATATCTCGCGCTTCCCTTTATGTTCC---	1994	CT
QY	725	rAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyrGlyGlyTyrLeuSe	745	
DB	1995	GAGTTTCGTGGACTCGCAAAAGATTTCCTTATTTGGTTGGTCGTATGGGGATATTTGAC	2054	
QY	745	rLeuMetAlaLeu---MetGlnArgSerAspIlePheArgValAlaIleAlaGlyAlaPr	764	
DB	2055	TCGTGAAGACCTTGGAAAGGACGAGGAGAACACTTTTAAAGTACGGGATGCTGTGTGCTC	2114	
QY	764	oValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMetGlyHisProAs	784	
DB	2115	GGTACCGATCGAGATTCTACGACTCTGTCTATACCGAAAGATACATGCACACTCTCA	2174	
QY	784	pGlnAsnGluGlnGlyTyrTyrLeuGlySerVal-----AlaMetGlnAlaGl	800	
DB	2175	AGAAAACCTTTCAGCGGTACGTGGAATCAAGTGTCCACATGTCACAGCGTTG-----	2226	

us-10-825-632-1.p2n.rnpbn

Tue Apr 18 08:18:33 2006

Qy 800 ulysPheProSerGluProAsnAtcLeuLeuLeuHisGlyPheLeuAspGluAsnVa 820
 Db 2227 -----GCCCAAGTAAACAGGTTCTACTAATGATGTTACAGGGATGACATGT 2276
 Qy 820 lHisPheAlaHisThrSerIleLeuLeuSerPheLeu---ValArgAlaGlyLysProTy 839
 Db 2277 TCACITCCAAATTCCTGAAATTTCTAGACCTTTTGAACCTTGAACAATGTCGAAATTA 2336
 Qy 839 rAspLeuGlnIleTyProGlnGluArgHisSerIleArg 852
 Db 2337 CGATGTCATGATTCCCTGATTCTGACCACAGTATAGA 2376

Search completed: April 15, 2006, 01:49:57
 Job time : 930 secs

CTGTTGATGTTCCCTGATTCTGACCACAGTATAGA

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 14, 2006, 19:35:26 ; Search time 15351 Seconds
(without alignments)
11553.095 Million cell updates/sec

Title: US-10-825-632-2
Perfect score: 3120
Sequence: 1 aagtgctaaagctccgagg.....agaattactaaaaaaaaa 3120

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_in.*
- 3: gb_env.*
- 4: gb_on.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pr.*
- 9: gb_ro.*
- 10: gb_sts.*
- 11: gb_sy.*
- 12: gb_un.*
- 13: gb_vi.*
- 14: gb_htg.*
- 15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3120	100.0	3120	6 AR651452	AR651452 Sequence
2	3120	100.0	3127	8 AF221634	AF221634 Homo sapi
3	3106.4	99.6	4829	6 AR631283	AR631283 Sequence
4	3106.4	99.6	4829	6 AX608735	AX608735 Sequence
5	3105.2	99.5	3143	6 AX354793	AX354793 Sequence
6	3101.8	99.4	3106	6 AX342633	AX342633 Sequence
7	2915.2	93.4	4535	8 BC040203	BC040203 Homo sapi
8	2824.4	90.5	4685	6 AR631288	AR631288 Sequence
9	2824.4	90.5	4685	6 AX608745	AX608745 Sequence
10	2806.4	89.9	4676	6 AR631287	AR631287 Sequence
11	2806.4	89.9	4676	6 AX608743	AX608743 Sequence
12	2805.4	89.9	3030	8 AY354202	AY354202 Homo sapi
13	2759.4	88.4	2797	6 AR48400	AR48400 Sequence
14	2656.4	85.1	3125	8 BC030688	BC030688 Homo sapi
15	2645.8	84.8	2649	6 AX354795	AX354795 Sequence
16	2644.4	84.8	2656	8 AY172659	AY172659 Homo sapi
17	2644.4	84.8	2671	6 AR631278	AR631278 Sequence
18	2644.4	84.8	2671	6 AX608725	AX608725 Sequence

19	2500.4	80.1	4523	6 AR631281	AR631281 Sequence
20	2500.4	80.1	4523	6 AX608731	AX608731 Sequence
21	2490	79.8	2842	6 AX405770	AX405770 Sequence
22	2445.4	78.4	2778	8 AK000290	AK000290 Homo sapi
23	2273.4	72.9	4799	9 BC043124	BC043124 Mus muscu
24	2273.4	72.9	4799	9 BC059222	BC059222 Mus muscu
25	2208	70.8	2510	6 AX338497	AX338497 Sequence
26	2132	68.3	2668	6 AX405771	AX405771 Sequence
27	2056.4	65.9	4309	6 AR631284	AR631284 Sequence
28	2056.4	65.9	4309	6 AX608737	AX608737 Sequence
29	1994.8	63.9	2161	6 BD157001	BD157001 Primer fo
30	1994.8	63.9	2161	6 AX878058	AX878058 Sequence
31	1994.8	63.9	2161	8 AK027826	AK027826 Homo sapi
32	1369	43.9	18341	14 AC161443	AC161443 Mus muscu
33	1347	43.2	1669	6 AR651454	AR651454 Sequence
34	1347	43.2	1669	8 AF221636	AF221636 Homo sapi
35	1330.8	42.7	2098	6 CQ721272	CQ721272 Sequence
36	1036	33.2	1197	6 AR651453	AR651453 Sequence
37	1036	33.2	1197	8 AF221635	AF221635 Homo sapi
c 38	967.4	31.0	194188	9 AC163387	AC163387 Mus muscu
39	926.4	29.7	1356	6 AR631282	AR631282 Sequence
40	926.4	29.7	1356	6 AX608733	AX608733 Sequence
41	901.6	28.9	3383	5 BC071112	BC071112 Xenopus l
42	882.2	28.3	3399	9 BC057631	BC057631 Mus muscu
43	862.6	27.6	2660	6 AX921883	AX921883 Sequence
44	862.6	27.6	2660	6 AX921885	AX921885 Sequence
45	862.2	27.6	1449	5 CR353491	CR353491 Gallus ga

ALIGNMENTS

RESULT 1
AR651452
LOCUS AR651452 3120 bp DNA linear PAT 20-APR-2005
DEFINITION Sequence, 2 from Patent US 6881564.
ACCESSION AR651452
VERSION AR651452.1 GI:62795938
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.

REFERENCE
1 (bases 1 to 3120)
AUTHORS Abbott, C.A. and Gorrell, M.D.
TITLE Dipeptidyl peptidases
JOURNAL Patent: US 6881564-A 2 19-APR-2005;
The University of Sydney, Sydney;
AUX;

FEATURES
source Location/Qualifiers
1..3120
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match	100.0%;	Score 3120;	DB 6;	Length 3120;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 3120;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	AAGTGCTAAAGCCTCCGAGCCCAAGCGCGTCTACTGCGCGCGTCTTCTTAGTCCG	60	
Db	1	AAGTGCTAAAGCCTCCGAGCCCAAGCGCGTCTACTGCGCGCGTCTTCTTAGTCCG	60	
Qy	61	CGTTCCGCGCGTGGTGTGTCAACCGCGCGCGCGCGAGAGAGCCACTGCAACACGAGACCG	120	
Db	61	CGTTCCGCGCGTGGTGTGTGTCAACCGCGCGCGCGAGAGAGCCACTGCAACACGAGACCG	120	
Qy	121	GAGTGGAGGCGCGCGAGCATGAAGCGCGCGCGCGTCTCCATAGCGCACCTCGGAGCGG	180	
Db	121	GAGTGGAGGCGCGCGAGCATGAAGCGCGCGCGCGTCTCCATAGCGCACCTCGGAGCGG	180	
Qy	181	TCCGGCGCGCGCGCGGGGGAGGAAATTCATATGCGCGAGCGAGCAATCGAAGCAACAG	240	
Db	181	TCCGGCGCGCGCGCGGGGGAGGAAATTCATATGCGCGAGCGAGCAATCGAAGCAACAG	240	

QY	241	CTGGGTGTGAGATATTTGAAACTGGGACTGTGAGGAAGATATTTGAATACAGAGTCGG	300	QY	1321	GAGGGAATATGCTTGGTCCATCTCTAGATCGCTCCAGACTCGCTACAGATAGTG	1380
Db	241	CTGGGTGTGAGATATTTGAAACTGGGACTGTGAGGAAGATATTTGAATACAGAGTCGG	300	Db	1321	GAGGGAATATGCTTGGTCCATCTCTAGATCGCTCCAGACTCGCTACAGATAGTG	1380
QY	301	CCTAAATTTGAGCCCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGCTGCTT	360	QY	1381	TTGATCTCAGCTGAATTTATTTATCCAGTAGAAGATGATGTTATGGAAGGAGAGATCTC	1440
Db	301	CCTAAATTTGAGCCCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGCTGCTT	360	Db	1381	TTGATCTCAGCTGAATTTATTTATCCAGTAGAAGATGATGTTATGGAAGGAGAGATCTC	1440
QY	361	GCCGATACAGAAATATCATGGCTACATGATGGCTAAGGCCACCATGATTTTCATGTTT	420	QY	1441	ATTGAGTCAGTCCCTGATCTCTGTGAGCCCACTAATTTCTATGAGAAACACACACATC	1500
Db	361	GCCGATACAGAAATATCATGGCTACATGATGGCTAAGGCCACCATGATTTTCATGTTT	420	Db	1441	ATTGAGTCAGTCCCTGATCTCTGTGAGCCCACTAATTTCTATGAGAAACACACACATC	1500
QY	421	GTGAGAGGATGATCCAGATGGACCTCATTTAGAGAGAAATCTTATACCTTCCCATGCT	480	QY	1501	TGGATAAATATCCATGACATCTTTTCATGTTTTTCCCAAGTCAAGAGGAAATTTGAG	1560
Db	421	GTGAGAGGATGATCCAGATGGACCTCATTTAGAGAGAAATCTTATACCTTCCCATGCT	480	Db	1501	TGGATAAATATCCATGACATCTTTTCATGTTTTTCCCAAGTCAAGAGGAAATTTGAG	1560
QY	481	GTTGAGACAGAGAAATACACTGTTTTTATCTGAAATCCCAAACTATCAATAGAGCA	540	QY	1561	TTTTTTTTTGGCTCTCGAATGCAAAACAGGTTTCCGTCATTATACAAAATTCATCTATT	1620
Db	481	GTTGAGACAGAGAAATACACTGTTTTTATCTGAAATCCCAAACTATCAATAGAGCA	540	Db	1561	TTTTTTTTTGGCTCTCGAATGCAAAACAGGTTTCCGTCATTATACAAAATTCATCTATT	1620
QY	541	GCAGTCTTAATGCTCTCTTGGAGCCCTTTTGGATCTTTTTCAGGCAACACTGGACTAT	600	QY	1621	TTAAAGGAAAGCAAAATATAAAGTCCAGTGGTGGCTGCCCTCCAAAGTATTTCAAG	1680
Db	541	GCAGTCTTAATGCTCTCTTGGAGCCCTTTTGGATCTTTTTCAGGCAACACTGGACTAT	600	Db	1621	TTAAAGGAAAGCAAAATATAAAGTCCAGTGGTGGCTGCCCTCCAAAGTATTTCAAG	1680
QY	601	GGAATGTATTCGAGAGAGAAAGAACTATTAAGAGAAAGAAACGCATTCGAAACAGTCGGA	660	QY	1681	TGTCTCTATCAAAGAGGAGATAGCAATTTACAGTGGTGAATGGGAAGTTCTTGGCCGGCAT	1740
Db	601	GGAATGTATTCGAGAGAGAAAGAACTATTAAGAGAAAGAAACGCATTCGAAACAGTCGGA	660	Db	1681	TGTCTCTATCAAAGAGGAGATAGCAATTTACAGTGGTGAATGGGAAGTTCTTGGCCGGCAT	1740
QY	661	ATTGCTTCTTACGATATTACCAAGGAAGTGGAACTTTCTGTTTCAAGCCGGTAGTGA	720	QY	1741	GGATCTAATATCCAAGTTTGATGAAGTCAAGAGGCTGGTATATTTTGAAGGCCACCAAGAC	1800
Db	661	ATTGCTTCTTACGATATTACCAAGGAAGTGGAACTTTCTGTTTCAAGCCGGTAGTGA	720	Db	1741	GGATCTAATATCCAAGTTTGATGAAGTCAAGAGGCTGGTATATTTTGAAGGCCACCAAGAC	1800
QY	721	ATTTATCAGTAAAGATGGAGGCCCAAGGATTTTACGCAACAACTTTTAAAGGCCCAAT	780	QY	1801	TGCCCTTTAGAGCATCACTGTAGTAGTCAAGTTCAGTAAATCCTCGAGAGGTGACAAGG	1860
Db	721	ATTTATCAGTAAAGATGGAGGCCCAAGGATTTTACGCAACAACTTTTAAAGGCCCAAT	780	Db	1801	TGCCCTTTAGAGCATCACTGTAGTAGTCAAGTTCAGTAAATCCTCGAGAGGTGACAAGG	1860
QY	781	CTAGTGGAACTAGTTGTCACCAACATACGATCGATCCAAATTTATGCCCGCTGATCCA	840	QY	1861	CTGACTGACCGTGGCTACTCACTTTCTGTCATCAGTCAAGCTGTGACTTCTTTATA	1920
Db	781	CTAGTGGAACTAGTTGTCACCAACATACGATCGATCCAAATTTATGCCCGCTGATCCA	840	Db	1861	CTGACTGACCGTGGCTACTCACTTTCTGTCATCAGTCAAGCTGTGACTTCTTTATA	1920
QY	841	GACTGATGCTTTTATACATAGCAACGATATTTGGATATCTAACATCGTAAACAGAGAA	900	QY	1921	AGTAAAGTAAAGTAAACAGAGAAATTTTGGGCCACCAATTTTGGATTCAGCAGGT	2040
Db	841	GACTGATGCTTTTATACATAGCAACGATATTTGGATATCTAACATCGTAAACAGAGAA	900	Db	1921	AGTAAAGTAAAGTAAACAGAGAAATTTTGGGCCACCAATTTTGGATTCAGCAGGT	2040
QY	901	GAAAGAGACTCACTTATGTCACAACTAGCTAGCCACATGGAAGAGATGCCAGATCA	960	QY	2041	CCTCTTCTGACTATACCTCCAGAAATTTTCTCTTTGAAAGTACTACTGGATTTACA	2100
Db	901	GAAAGAGACTCACTTATGTCACAACTAGCTAGCCACATGGAAGAGATGCCAGATCA	960	Db	2041	CCTCTTCTGACTATACCTCCAGAAATTTTCTCTTTGAAAGTACTACTGGATTTACA	2100
QY	961	GCTGGAGTCGCTACCTTTGTTCTCCAAGAGAAATTTGATATATTCGCTATTCGTTGG	1020	QY	2101	TTGATGCGGATGCTCTACAAAGCTCATGATCTACAGGCTCGAAAGAAATATCCTACTGTG	2160
Db	961	GCTGGAGTCGCTACCTTTGTTCTCCAAGAGAAATTTGATATATTCGCTATTCGTTGG	1020	Db	2101	TTGATGCGGATGCTCTACAAAGCTCATGATCTACAGGCTCGAAAGAAATATCCTACTGTG	2160
QY	1021	TGTCCAAAGCTGAAACAACTCCAGTGGTAAATTTCTAGAAATTCATATGAAGAA	1080	QY	2161	CTGTTTCATATGTTGGTCTCTCAGTGCAGTTGGTGAATATCGGTTTAAAGGAGTCAAG	2220
Db	1021	TGTCCAAAGCTGAAACAACTCCAGTGGTAAATTTCTAGAAATTCATATGAAGAA	1080	Db	2161	CTGTTTCATATGTTGGTCTCTCAGTGCAGTTGGTGAATATCGGTTTAAAGGAGTCAAG	2220
QY	1081	AATGATGAATCTGAGGTGGAATTTATTCATGTTACATCCCTATGTTGGAAACAGAGAG	1140	QY	2221	TATTTCCGCTTGAATACCCCTAGCCCTCTCAGGTTATGTTGGTGTAGTAGAGACAAAGG	2280
Db	1081	AATGATGAATCTGAGGTGGAATTTATTCATGTTACATCCCTATGTTGGAAACAGAGAG	1140	Db	2221	TATTTCCGCTTGAATACCCCTAGCCCTCTCAGGTTATGTTGGTGTAGTAGAGACAAAGG	2280
QY	1141	GCAGATTCATTCGCTTATCCTAAACAGGTAAGCAATCCCTAAAGTCACTTTTAAAGATG	1200	QY	2281	GGATCTGTCCCGGGCTTAAATTTGAAAGCCCTTTAAATATAAAATGGGTCAATAA	2340
Db	1141	GCAGATTCATTCGCTTATCCTAAACAGGTAAGCAATCCCTAAAGTCACTTTTAAAGATG	1200	Db	2281	GGATCTGTCCCGGGCTTAAATTTGAAAGCCCTTTAAATATAAAATGGGTCAATAA	2340
QY	1201	TCAGAAATATGATGATGCTGAAGGAGGATCATAGATGTCATAGATAAGGAACTAAT	1260	QY	2341	GAAATGGAACGATCAGGTGGAAGACTCCCAATATCTAGCTTCCGATATGATTTTCATTGAC	2400
Db	1201	TCAGAAATATGATGATGCTGAAGGAGGATCATAGATGTCATAGATAAGGAACTAAT	1260	Db	2341	GAAATGGAACGATCAGGTGGAAGACTCCCAATATCTAGCTTCCGATATGATTTTCATTGAC	2400
QY	1261	CAACCTTTGAGATCTATTTGAAGGAGTTGAATATATTGCCAGAGCTGGATGACTCCT	1320	QY	2401	TTAGATCGTGTGGGCATCCACGGCTGGTCTCTATGGAGGATACCTCTCCTCGATGGCATTA	2460
Db	1261	CAACCTTTGAGATCTATTTGAAGGAGTTGAATATATTGCCAGAGCTGGATGACTCCT	1320	Db	2401	TTAGATCGTGTGGGCATCCACGGCTGGTCTCTATGGAGGATACCTCTCCTCGATGGCATTA	2460

FEATURES		Location/Qualifiers	
source			
Db	2401	TTAGATCGTGTGGGCATCCACGGCTGGTCTCTATGGAGGATACCTCTCCCTGATGGCATTA	2460
Qy	2461	ATGCAGAGGTTCAGATATCTTTCAGGGTGTCTATTGCTGGGGCCCGAGTCACCTCTGTGGATC	2520
Db	2461	ATGCAGAGGTTCAGATATCTTTCAGGGTGTCTATTGCTGGGGCCCGAGTCACCTCTGTGGATC	2520
Qy	2521	TTCTATGATACAGGATACACGGAACGTTATATGGGTACCCCTGACCAAGATGAACAGGGC	2580
Db	2521	TTCTATGATACAGGATACACGGAACGTTATATGGGTACCCCTGACCAAGATGAACAGGGC	2580
Qy	2581	TATTACTTAGGATCTGTGGCCATGCAAGCAAGAAAGTTCCCTCTGACCAAAATCGTTTA	2640
Db	2581	TATTACTTAGGATCTGTGGCCATGCAAGCAAGAAAGTTCCCTCTGACCAAAATCGTTTA	2640
Qy	2641	CTGCTCTTACATGTTTCTCGATGAGATGCAATTTTGGACATACACAGTATATTACTG	2700
Db	2641	CTGCTCTTACATGTTTCTCGATGAGATGCAATTTTGGACATACACAGTATATTACTG	2700
Qy	2701	AGTTTTTTTACTGAGGGCTGGAAAGCCATATGATTTTACAGATCTATCTCTCAGGAGACAC	2760
Db	2701	AGTTTTTTTACTGAGGGCTGGAAAGCCATATGATTTTACAGATCTATCTCTCAGGAGACAC	2760
Qy	2761	AGCATAGAGTTCCTGAATCGGAGAACATATGAACTGCAATCTTTTGGCACTACTTCAA	2820
Db	2761	AGCATAGAGTTCCTGAATCGGAGAACATATGAACTGCAATCTTTTGGCACTACTTCAA	2820
Qy	2821	GAACACCTTGGATCAGTATCTGCTCTTAAAGTGATATAATTTGACCTGTGAGAAC	2880
Db	2821	GAACACCTTGGATCAGTATCTGCTCTTAAAGTGATATAATTTGACCTGTGAGAAC	2880
Qy	2881	TCTCTGTATACACTGGCTATTTAACCAATCAGAGGTTTAAATCAACAGAAACACAGA	2940
Db	2881	TCTCTGTATACACTGGCTATTTAACCAATCAGAGGTTTAAATCAACAGAAACACAGA	2940
Qy	2941	ATTGATCATCACAATTTTGATPACTGCGCATGTAACTACTCTCTGAAATAAATGTGGTG	3000
Db	2941	ATTGATCATCACAATTTTGATPACTGCGCATGTAACTACTCTCTGAAATAAATGTGGTG	3000
Qy	3001	CCATGAGGGGTACGGTTTGGTAGTAATCTAATACCTTAACCCACATGCTCAAAA	3060
Db	3001	CCATGAGGGGTACGGTTTGGTAGTAATCTAATACCTTAACCCACATGCTCAAAA	3060
Qy	3061	TCAAATGATACATATCTCTGAGAGCCAGCAATACCAATGAGATTTACTTAAAAAAAAA	3120
Db	3061	TCAAATGATACATATCTCTGAGAGCCAGCAATACCAATGAGATTTACTTAAAAAAAAA	3120
RESULT 2			
AF221634			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			

Qy	481	GGTGAGAACAGAGAAAATACACTGTTTATTCGTGAATTC	Db	1561	TTTATTTTGGCTCTGAATGCNAAACAGGTTTCCGTCATTTATACAAAATTACATCTATT
Db	481	GGTGAGAACAGAGAAAATACACTGTTTATTCGTGAATTC	Qy	1621	TTAAAGGAACAAATATATAACGATCCAGTGGTGGCTGCCTCCCTCAAGTGCATTTCAAG
Qy	541	GCAGTCTTAATGCTCTCTTGGAAGCCTCTTTTGGATCTTTT	Db	1621	TTAAAGGAACAAATATATAACGATCCAGTGGTGGCTGCCTCCCTCAAGTGCATTTCAAG
Db	541	GCAGTCTTAATGCTCTCTTGGAAGCCTCTTTTGGATCTTTT	Qy	1681	TGCTCTATCAAAAGAGAGATAGCAATTTACAGTGGTGAATGGGAAGTCTTGGCCGGCAT
Qy	601	GGAATGTATTCCTCGAGAAGAACTATTAAAGAGAAGAAAACGCAATTGGAACAGTCGGA	Db	1681	TGCTCTATCAAAAGAGAGATAGCAATTTACAGTGGTGAATGGGAAGTCTTGGCCGGCAT
Db	601	GGAATGTATTCCTCGAGAAGAACTATTAAAGAGAAGAAAACGCAATTGGAACAGTCGGA	Qy	1741	GGATCTAATATCCAAAGTTGATGAAGTCCAGAGCTGGTATATTTTGAAGGACCAAGAC
Qy	661	ATTGCTCTTACGATTAATCAACAGGAAGTGAACATTTCTGTTTCAAGCCGGTAGTGA	Db	1741	GGATCTAATATCCAAAGTTGATGAAGTCCAGAGCTGGTATATTTTGAAGGACCAAGAC
Db	661	ATTGCTCTTACGATTAATCAACAGGAAGTGAACATTTCTGTTTCAAGCCGGTAGTGA	Qy	1801	TCCCTTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAAATCCTGGAGAGTGACAAAG
Qy	721	ATTATCAGTAAAGATGGAGGCCACAGGATTTACGCAACAACTTTTAAGGCCCAAT	Db	1801	TCCCTTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAAATCCTGGAGAGTGACAAAG
Db	721	ATTATCAGTAAAGATGGAGGCCACAGGATTTACGCAACAACTTTTAAGGCCCAAT	Qy	1861	CTGACTGACCCGTGGCTACTCACTTCTGTGTCAGTCAGTCAGCACTGTGACTTCTTTATA
Qy	781	CTAGTGGAACTAGTTGTCCCAACATACGGAATGGATCCAAATTTAGCCGCTGATCCA	Db	1861	CTGACTGACCCGTGGCTACTCACTTCTGTGTCAGTCAGTCAGCACTGTGACTTCTTTATA
Db	781	CTAGTGGAACTAGTTGTCCCAACATACGGAATGGATCCAAATTTAGCCGCTGATCCA	Qy	1921	AGTAAAGTATAGTAAACCAAGAAATCCACACTGTGTCTTCAAGCTATCAAGTCTCT
Qy	841	GACTGGATGCTTTTATACATAGCAACGATATTTTGGATATCTTAAACATGCTAACCAAGNA	Db	1921	AGTAAAGTATAGTAAACCAAGAAATCCACACTGTGTCTTCAAGCTATCAAGTCTCT
Db	841	GACTGGATGCTTTTATACATAGCAACGATATTTTGGATATCTTAAACATGCTAACCAAGNA	Qy	1981	GAAGATGCCCAACTTTCGCAAAACAAAGAAATTTTGGGCCACCAATTTTGGATTTCAGCAGGT
Qy	901	GAAAGAGACTCACTTATGTGCACATAGCTAGCCCAACATGGAAGAAGATGCCAGATCA	Db	1981	GAAGATGCCCAACTTTCGCAAAACAAAGAAATTTTGGGCCACCAATTTTGGATTTCAGCAGGT
Db	901	GAAAGAGACTCACTTATGTGCACATAGCTAGCCCAACATGGAAGAAGATGCCAGATCA	Qy	2041	CCTCTTCGTACTATCTCTCCAGAAATTTCTCTTTGAAAGTACTACTGATTTTACA
Qy	961	GCTGGAGTGGCTACCTTGTCTTCCAGGAAGAAATTTGATAGATATTTCTGGCTATTTGGTG	Db	2041	CCTCTTCGTACTATCTCTCCAGAAATTTCTCTTTGAAAGTACTACTGATTTTACA
Db	961	GCTGGAGTGGCTACCTTGTCTTCCAGGAAGAAATTTGATAGATATTTCTGGCTATTTGGTG	Qy	2101	TTGTATGGATGCTCTCAAGCCTCATGATCTACAGCTCAGCTCGGAAGAATATCTACTGTG
Qy	1021	TGTCCAAAGCTGAAACAACTCCAGTGGTGGTAAATTTCTTGAATTTCTATATGAAGAA	Db	2101	TTGTATGGATGCTCTCAAGCCTCATGATCTACAGCTCAGCTCGGAAGAATATCTACTGTG
Db	1021	TGTCCAAAGCTGAAACAACTCCAGTGGTGGTAAATTTCTTGAATTTCTATATGAAGAA	Qy	2161	CTGTTTCATATATGGTGGTCTCAGTGCAATGGTGAATTAATCGGTTTAAAGAGTCAAG
Qy	1081	AATGATGAATCTGAGTGGAAATTTATCATGTACATCCCTATGTTGGAAACAAAGGAG	Db	2161	CTGTTTCATATATGGTGGTCTCAGTGCAATGGTGAATTAATCGGTTTAAAGAGTCAAG
Db	1081	AATGATGAATCTGAGTGGAAATTTATCATGTACATCCCTATGTTGGAAACAAAGGAG	Qy	2221	TATTTCCGCTTGAATACCCCTAGCCTCTAGGTTATGTGGTTAGTGTAGTAGACAAACAG
Qy	1141	GCAGATTCATTCGGTTATCTTAAACAGGTACAGCAATCTTAAAGTCACTTTTAAAGATG	Db	2221	TATTTCCGCTTGAATACCCCTAGCCTCTAGGTTATGTGGTTAGTGTAGTAGACAAACAG
Db	1141	GCAGATTCATTCGGTTATCTTAAACAGGTACAGCAATCTTAAAGTCACTTTTAAAGATG	Qy	2281	GGATCCTGTCAACGAGGGCTTAAATTTGAAGCGCCTTTAAATATAAAATGGGTCAAATA
Qy	1201	TCAGAAATATATGATGCTGAGGAAGGATCATAGTGTCTAGATAGGAACATAAT	Db	2281	GGATCCTGTCAACGAGGGCTTAAATTTGAAGCGCCTTTAAATATAAAATGGGTCAAATA
Db	1201	TCAGAAATATATGATGCTGAGGAAGGATCATAGTGTCTAGATAGGAACATAAT	Qy	2341	GAAATTTGACGATCAGGTGGAAGGACTCCCAATATCTAGCTTCTCGATATGATTTCAATTGAC
Qy	1261	CAACCTTTGAGATTCATTTTGAAGAGTTGATATATGTCAGAGCTGATGACTCCT	Db	2341	GAAATTTGACGATCAGGTGGAAGGACTCCCAATATCTAGCTTCTCGATATGATTTCAATTGAC
Db	1261	CAACCTTTGAGATTCATTTTGAAGAGTTGATATATGTCAGAGCTGATGACTCCT	Qy	2401	TTAGATCGTGTGGGCATCCACGGCTGGTCTTATGGAGGATACCTCTCCCTGATGGCATTA
Qy	1321	GAGGAAATATGCTTGGTCCATCTCTACATGAGTCCGCTCCAGACTCGCTACAGATAGT	Db	2401	TTAGATCGTGTGGGCATCCACGGCTGGTCTTATGGAGGATACCTCTCCCTGATGGCATTA
Db	1321	GAGGAAATATGCTTGGTCCATCTCTACATGAGTCCGCTCCAGACTCGCTACAGATAGT	Qy	2461	ATCAGAGGTTCAGATATCTTCAAGGTTGCTATTTGCTGGGGCCCGAGTCACTCTCTGTGATC
Qy	1381	TTGATCTCACCTGAAATTTATTTATCCAGTGAAGATGATGTTATGGAAAGCGAGACTC	Db	2461	ATCAGAGGTTCAGATATCTTCAAGGTTGCTATTTGCTGGGGCCCGAGTCACTCTCTGTGATC
Db	1381	TTGATCTCACCTGAAATTTATTTATCCAGTGAAGATGATGTTATGGAAAGCGAGACTC	Qy	2521	TTCTATGATACAGGATACCGGAACTTATATGGGTCACTTACCCCTGACCAAGATGAACAGGC
Qy	1441	ATTGAGTCAGTGGCTGATCTGTGACGCCACTTAAATATCTATGAAGAAACAAACAGATC	Db	2521	TTCTATGATACAGGATACCGGAACTTATATGGGTCACTTACCCCTGACCAAGATGAACAGGC
Db	1441	ATTGAGTCAGTGGCTGATCTGTGACGCCACTTAAATATCTATGAAGAAACAAACAGATC	Qy	2581	TATTAATTTAGGATCTGTGGCCATGCAAGCGAAAGTTCCCTCTGACCAAAATCGTTTA
Qy	1501	TGGATAAATATCAATGACATCTTTCAATGTTTTTCCCAAGTCAACGAGGAAATTTGAG	Db	2581	TATTAATTTAGGATCTGTGGCCATGCAAGCGAAAGTTCCCTCTGACCAAAATCGTTTA
Db	1501	TGGATAAATATCAATGACATCTTTCAATGTTTTTCCCAAGTCAACGAGGAAATTTGAG	Qy	2641	CTGCTCTTACATCGTTTCTTGGATGGAATGTCATTTTGGACATACCAAGTATATTACTG
Qy	1561	TTTATTTTGGCTCTGAATGCNAAACAGGTTTCCGTCATTTATACAAAATTACATCTATT			

Db	2641	CTGCTCTTACATGGTTTCTCTGGATGAGAAATGTCCATTTTGGCACATACCACTATATTACTG	2700
Qy	2701	AGTTTTTTAGTCAGGCGCTGGAAGCCATATGATTTTACAGATCTATCTCTCAGGAGAGACAC	2760
Db	2701	AGTTTTTTAGTCAGGCGCTGGAAGCCATATGATTTTACAGATCTATCTCTCAGGAGAGACAC	2760
Qy	2761	AGCATAAAGAGTTCTCGAATCGGGAGAACATTTATGAATCGCATCTTTTGGCACTACCTTCAA	2820
Db	2761	AGCATAAAGAGTTCTCGAATCGGGAGAACATTTATGAATCGCATCTTTTGGCACTACCTTCAA	2820
Qy	2821	GAAAACTTTGGATCAGATTTGCTGCTCTTAAAGTGATATATTTTGGACCTGTGTAGAAC	2880
Db	2821	GAAAACTTTGGATCAGATTTGCTGCTCTTAAAGTGATATATTTTGGACCTGTGTAGAAC	2880
Qy	2881	TCTCTGGTATACACTGGCTATTTAAACCAATGAGAGGTTTAAATCAACAGAAAAACACAGA	2940
Db	2881	TCTCTGGTATACACTGGCTATTTAAACCAATGAGAGGTTTAAATCAACAGAAAAACACAGA	2940
Qy	2941	ATTGATCATCACATTTTGATACCTGCCATGTAAACATCTACCTCGAAAAATTAATGTGGTG	3000
Db	2941	ATTGATCATCACATTTTGATACCTGCCATGTAAACATCTACCTCGAAAAATTAATGTGGTG	3000
Qy	3001	CCATGAGGGGCTACAGGTTTGTGTAGTAACTTAATACCTTAACCCCATGCTCAAAA	3060
Db	3001	CCATGAGGGGCTACAGGTTTGTGTAGTAACTTAATACCTTAACCCCATGCTCAAAA	3060
Qy	3061	TCAAATGATACATATTCCTGAGAGACCCAGCAATACCAATAGAAATTAATTAAGAAAAA	3120
Db	3061	TCAAATGATACATATTCCTGAGAGACCCAGCAATACCAATAGAAATTAATTAAGAAAAA	3120

RESULT 3

AR631283

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

ORIGIN

Query Match

Best Local Similarity

Matches 3119; Conservative

0; Mismatches

1; Indels

2; Gaps

1;

Qy	1	AAGTGTAAAGCTTCGAGGCCAAGCCGCTGCTACTGCGCGCGCTGCTTCTTAGTCCG	60
Db	1	AAGTGTAAAGCTTCGAGGCCAAGCCGCTGCTACTGCGCGCGCTGCTTCTTAGTCCG	60
Qy	61	CGTTCCGCGCTGGTGTGTCACCGCGCGCGCGCGAGAGCCACTGCAACACGAGACCG	120
Db	61	CGTTCCGCGCTGGTGTGTCACCGCGCGCGCGCGAGAGCCACTGCAACACGAGACCG	120
Qy	121	GAGTGGAGGGCGCGCAGCATGAAGCGCGCGCGCGCTCCATAGCGCAGCTCGGGACGG	180
Db	121	GAGTGGAGGGCGCGCAGCATGAAGCGCGCGCGCGCTCCATAGCGCAGCTCGGGACGG	180
Qy	181	TCCGGGCGGGCGGGGAGGAAAATGCAATGCAATGCGGAGCAGCAATGGAACAGAACAG	240
Db	181	TCCGGGCGGGCGGGGAGGAAAATGCAATGCAATGCGGAGCAGCAATGGAACAGAACAG	240

Db	1321	GAGGAAATATGCTTGGTGCATCTCTACATAGATCGCTCCACACTCGCTCAGATAGTG	1380
Qy	1381	TTGATCTCACCTGAAATATTTATCCAGTAGAAGATGATGTTATGGAAAGCGAGACTC	1440
Db	1381	TTGATCTCACCTGAAATATTTATCCAGTAGAAGATGATGTTATGGAAAGCGAGACTC	1440
Qy	1441	ATTGAGTCAGTGCCTGATCTCTGTGAAGCCACTAAATTTATCTATGAAGAAACACAGACATC	1500
Db	1441	ATTGAGTCAGTGCCTGATCTCTGTGAAGCCACTAAATTTATCTATGAAGAAACACAGACATC	1500
Qy	1501	TGGATAAATATCCATGACATCTTTCATGTTTTTCCCAAGTCAAGAGAGGAAATTCAG	1560
Db	1501	TGGATAAATATCCATGACATCTTTCATGTTTTTCCCAAGTCAAGAGAGGAAATTCAG	1560
Qy	1561	TTTATTTTGGCTCTGATGCAACACAGGTTTCGGTCAATTTATACAAAATACATCTATT	1620
Db	1561	TTTATTTTGGCTCTGATGCAACACAGGTTTCGGTCAATTTATACAAAATACATCTATT	1620
Qy	1621	TTAAAGGAAAGCAATATTAACGATCCAGTGGTGGCTGCTCCAGTGCATTTCAAG	1680
Db	1621	TTAAAGGAAAGCAATATTAACGATCCAGTGGTGGCTGCTCCAGTGCATTTCAAG	1680
Qy	1681	TGTCTCTATCAAGAGGAGATAGCAATTTACAGTGGTGGATGGGAAGTTCTTGGCGGCAT	1740
Db	1681	TGTCTCTATCAAGAGGAGATAGCAATTTACAGTGGTGGATGGGAAGTTCTTGGCGGCAT	1740
Qy	1741	GGATCTAATATCAAGTTGATGAAGTCAGAGGCTGGTATPATTTTGAAGGCAACCAAGAC	1800
Db	1741	GGATCTAATATCAAGTTGATGAAGTCAGAGGCTGGTATPATTTTGAAGGCAACCAAGAC	1800
Qy	1801	TCCCTTTTAGAGCATCACCTGTAGTACGTACGTAAATCCTGGAGAGGTGACAAGG	1860
Db	1801	TCCCTTTTAGAGCATCACCTGTAGTACGTACGTAAATCCTGGAGAGGTGACAAGG	1860
Qy	1861	CTGACTGACCGTGGCTACTCACATTTCTGTGCACTCAGTCAGCACTGTGACTTTTATA	1920
Db	1861	CTGACTGACCGTGGCTACTCACATTTCTGTGCACTCAGTCAGCACTGTGACTTTTATA	1920
Qy	1921	AGTAACTAGTAAACAGAGAAATCCACACTGTGTCCCTTTTACAAAGCTATCAAGTCTCT	1980
Db	1921	AGTAACTAGTAAACAGAGAAATCCACACTGTGTCCCTTTTACAAAGCTATCAAGTCTCT	1980
Qy	1981	GAAGATGACCCCACTGTCGAAACAAAGGAAATTTTGGGCCACCATTTTGGATTCCAGCAGGT	2040
Db	1981	GAAGATGACCCCACTGTCGAAACAAAGGAAATTTTGGGCCACCATTTTGGATTCCAGCAGGT	2040
Qy	2041	CCTCTTCTGACTATCTCTCCAGAAATTTCTCTTTTGAAGTACTACTGGATTACA	2100
Db	2041	CCTCTTCTGACTATCTCTCCAGAAATTTCTCTTTTGAAGTACTACTGGATTACA	2100
Qy	2101	TTGTATGGATGCTCTACAGGCTCATGATCTACAGGCTGGAAAGAAATATCTACTGTG	2160
Db	2101	TTGTATGGATGCTCTACAGGCTCATGATCTACAGGCTGGAAAGAAATATCTACTGTG	2160
Qy	2161	CTGTTTCATATATGGTGG--TCCTCAGGTGCAAGTTGGTGAATTAATTCGGTTTAAAGGAGTCA	2218
Db	2161	CTGTTTCATATATGGTGGTCTCTCCTCAGGTGCAAGTTGGTGAATTAATTCGGTTTAAAGGAGTCA	2220
Qy	2219	AGTATTTCCGCTTGAATACCTTAGCTCTCTAGGTATGTTGGTTGTAGTGATAGACAACA	2278
Db	2221	AGTATTTCCGCTTGAATACCTTAGCTCTCTAGGTATGTTGGTTGTAGTGATAGACAACA	2280
Qy	2279	GGGATCCTGTGACCCAGGCTTAAATTTGAAGGCGCTTTTAAATATATAAAATGGGTCAA	2338
Db	2281	GGGATCCTGTGACCCAGGCTTAAATTTGAAGGCGCTTTTAAATATATAAAATGGGTCAA	2340
Qy	2339	TAGAATTTGACGATCAGGTGGAAGGACTCCATATCTAGCTTCTCGATATGATTTTCAATG	2398
Db	2341	TAGAATTTGACGATCAGGTGGAAGGACTCCATATCTAGCTTCTCGATATGATTTTCAATG	2400
Qy	2399	ACTTAGATCTGTGGGATCCAGGCTGGTCTTATGGAGGATACCTCTCCTCGATGGCAT	2458
Db	2401	ACTTAGATCTGTGGGATCCAGGCTGGTCTTATGGAGGATACCTCTCCTCGATGGCAT	2460
Qy	2459	TAATGTCAGAGTCAAGATATCTTCAGGGTTGCTATTCAGGGCCCGCAGTCACTCTGTGGA	2518
Db	2461	TAATGTCAGAGTCAAGATATCTTCAGGGTTGCTATTCAGGGCCCGCAGTCACTCTGTGGA	2520
Qy	2519	TCTTCTATGATACAGGATACACGGAAGTTATATGGTCACTCCCTGACCAAGATGAACAGG	2578
Db	2521	TCTTCTATGATACAGGATACACGGAAGTTATATGGTCACTCCCTGACCAAGATGAACAGG	2580
Qy	2579	GCTATTACTTAGATCTGTGGCCATGACGGAAGTTCCCTCTGGAACCAATCGTT	2638
Db	2581	GCTATTACTTAGATCTGTGGCCATGACGGAAGTTCCCTCTGGAACCAATCGTT	2640
Qy	2639	TACTGCTCTTATCATGGTTTCTCGATGAGAAATGTCCATTTTGGCAATACAGTATATTAC	2698
Db	2641	TACTGCTCTTATCATGGTTTCTCGATGAGAAATGTCCATTTTGGCAATACAGTATATTAC	2700
Qy	2699	TGAGTTTTTTTATGTCAGGCTGGAAAGCCATATGATTTTACAGATCTATTCCTCAGGAGAGAC	2758
Db	2701	TGAGTTTTTTTATGTCAGGCTGGAAAGCCATATGATTTTACAGATCTATTCCTCAGGAGAGAC	2760
Qy	2759	ACAGCATAAAGATTCTTGAATCGGGAGAAACATTTATGAATCTGCACTCTTTTGCACCTACCTTC	2818
Db	2761	ACAGCATAAAGATTCTTGAATCGGGAGAAACATTTATGAATCTGCACTCTTTTGCACCTACCTTC	2820
Qy	2819	AAGAAACCTTGGATCAAGTATTCCTCTCTTAAAGTGATATAATTTTGAACCTGTGTAGA	2878
Db	2821	AAGAAACCTTGGATCAAGTATTCCTCTCTTAAAGTGATATAATTTTGAACCTGTGTAGA	2880
Qy	2879	ACTCTCTGGTATACACTGGCTATTTAACCAAAATGAGGAGGTTTAAATCAACAGAAAAACACA	2938
Db	2881	ACTCTCTGGTATACACTGGCTATTTAACCAAAATGAGGAGGTTTAAATCAACAGAAAAACACA	2940
Qy	2939	GAATTGATCATCACTTTTGTATCTGCACTGCAATCTACTCTCTGAAATATAATGTGG	2998
Db	2941	GAATTGATCATCACTTTTGTATCTGCACTGCAATCTACTCTCTGAAATATAATGTGG	3000
Qy	2999	TGCCATGCAAGGCTCTACGGTTTGTGTAGTAAATCTAAATACCTTAAACCCACATGCTCAA	3058
Db	3001	TGCCATGCAAGGCTCTACGGTTTGTGTAGTAAATCTAAATACCTTAAACCCACATGCTCAA	3060
Qy	3059	AATCAAAATGATACATATTCCTGAGAGACCCAGCAATACCATAAGAAATTTCTTAAARAAA	3118
Db	3061	AATCAAAATGATACATATTCCTGAGAGACCCAGCAATACCATAAGAAATTTCTTAAARAAA	3120
Qy	3119	AA 3120	
Db	3121	AA 3122	
RESULT 4			
AX608735			
LOCUS			
Sequence 12 from Patent WO0231134.			
DEFINITION			
AX608735			
ACCESSION			
AX608735.1 GI:28404303			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
Homo sapiens (human)			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;			
Hominidae; Homo.			
REFERENCE			
1 .4829			
QI.S., Akinsanya, K.O., Riviere, P.J. and Junien, J.L.			
Novel serine protease genes related to dppiv			
Patent: WO 0231134-A 12 18-APR-2002;			
Ferring BV (NL)			
FEATURES			
source			
Location/Qualifiers			
1..4829			
/organism="Homo sapiens"			
/mol_type="unassigned DNA"			
/db_xref="taxon:9606"			

AX608735 4829 bp DNA linear PAT 17-FEB-2003

Sequence 12 from Patent WO0231134.

AX608735

AX608735.1 GI:28404303

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

1 .4829

QI.S., Akinsanya, K.O., Riviere, P.J. and Junien, J.L.

Novel serine protease genes related to dppiv

Patent: WO 0231134-A 12 18-APR-2002;

Ferring BV (NL)

FEATURES

source

Location/Qualifiers

1..4829

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

ORIGIN

Query Match	99.6%	Score 3106.4	DB 6	Length 4829
Best Local Similarity	99.9%	Pred. No. 0		
Matches 3119	Conservative	0	Mismatches 1	Indels 2

Qy	1	AAAGTCTAAAGCCTCCGAGGCCAAGCGCGTGTCTACTGCGCGCGCTGCTTCTTAGTGTCCG	60
Db	1	AAAGTCTAAAGCCTCCGAGGCCAAGCGCGTGTCTACTGCGCGCGCTGCTTCTTAGTGTCCG	60
Qy	61	CGTTCCGCGCCTGGTGTGTCAACGCGCGCGCGCGCGAGGAAGCCACTGCAACACGAGCACG	120
Db	61	CGTTCCGCGCCTGGTGTGTCAACGCGCGCGCGCGCGAGGAAGCCACTGCAACACGAGCACG	120
Qy	121	GAGTGAAGCGCGCGCAGCATGAAGCGCGCAGCGCCGCTTCATAGCGCATCGTGGGACGG	180
Db	121	GAGTGAAGCGCGCGCAGCATGAAGCGCGCAGCGCCGCTTCATAGCGCATCGTGGGACGG	180
Qy	181	TCCGGCGCGCGCGGGGGAAGGAATAATGCAACATGGCAGCAGCAATGGAAACAGACAG	240
Db	181	TCCGGCGCGCGCGGGGGAAGGAATAATGCAACATGGCAGCAGCAATGGAAACAGACAG	240
Qy	241	CTGGGTGTGGAGATATTTGAAACTGCGGCACTGTGAGGAGAAATTAATGAATCAACAGATCGG	300
Db	241	CTGGGTGTGGAGATATTTGAAACTGCGGCACTGTGAGGAGAAATTAATGAATCAACAGATCGG	300
Qy	301	CCTAAATTTGGAGCCTTTTATGTGAGCGGTATTTCTTGGAGTCAGCTTAAAAGCTGCTTT	360
Db	301	CCTAAATTTGGAGCCTTTTATGTGAGCGGTATTTCTTGGAGTCAGCTTAAAAGCTGCTTT	360
Qy	361	GCGGATACCCAGAAAAATATCATGGCTACATGATGGCTTAAGGCCACCAATGATTTCAATGTTT	420
Db	361	GCGGATACCAAGAAAAATATCATGGCTACATGATGGCTTAAGGCCACCAATGATTTCAATGTTT	420
Qy	421	GTGAAGAGGAATGATCCAGATGGACCTCAATTCAGACAGAAATCTATTACTCTTGGCAATGCT	480
Db	421	GTGAAGAGGAATGATCCAGATGGACCTCAATTCAGACAGAAATCTATTACTCTTGGCAATGCT	480
Qy	481	GGTGAGAACAGAAAAATACACTGTTTTATTCTGAAATTTCCCAAAACTATCAATAGACGA	540
Db	481	GGTGAGAACAGAAAAATACACTGTTTTATTCTGAAATTTCCCAAAACTATCAATAGACGA	540
Qy	541	CGAGTCTTAACTCTCTTTGGAAGCCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT	600
Db	541	CGAGTCTTAAATGCTCTCTTTGGAAGCCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT	600
Qy	601	GGAATGTATTTCTCGAAGAAAGAAACTATTAAAGAAAGAAAAACGCAATTTGGAACAGTCGGA	660
Db	601	GGAATGTATTTCTCGAAGAAAGAAACTATTAAAGAAAGAAAAACGCAATTTGGAACAGTCGGA	660
Qy	661	ATTGCTTCTTACGATTTATCCAGAGGAAGTGGAACTTTCTGTTTTCAAGCCGCTAGTGGGA	720
Db	661	ATTGCTTCTTACGATTTATCCAGAGGAAGTGGAACTTTCTGTTTTCAAGCCGCTAGTGGGA	720
Qy	721	ATTATTACGCTAAAAAGATGGAGGCCCAAGAGATTTTACGCAACAACCTTTTAAGGCCCAAT	780
Db	721	ATTATTACGCTAAAAGATGGAGGCCCAAGAGATTTTACGCAACAACCTTTTAAGGCCCAAT	780
Qy	781	CTAGTGGAAACTAGTGTGTTGCCAACAATACGATGGATCCAAAATTTATGCCCCGCTGATCCA	840
Db	781	CTAGTGGAAACTAGTGTGTTGCCAACAATACGATGGATCCAAAATTTATGCCCCGCTGATCCA	840
Qy	841	GACTCGATGCTTTTTATACATAGCAACGATATTTGGATATCTAAACATCTGAACACAGAGAA	900
Db	841	GACTCGATGCTTTTTATACATAGCAACGATATTTGGATATCTAAACATCTGAACACAGAGAA	900
Qy	901	GAAAGGAGACTCACTTATGTGCAACAATAGCTAGCCAAACATGGAAAGAGATGCCAGATCA	960
Db	901	GAAAGGAGACTCACTTATGTGCAACAATAGCTAGCCAAACATGGAAAGAGATGCCAGATCA	960
Qy	961	GCTGGAGTCGCTACCTTTGTTCTCAAGAGAAATTTGATAGATATTTCTGGCTATTTGGTG	1020
Db	961	GCTGGAGTCGCTACCTTTGTTCTCAAGAGAAATTTGATAGATATTTCTGGCTATTTGGTG	1020

Qy	1021	TGTC	CAAAAGCTG	GAAG	CAAA	CAAC	TCC	CAGTGGTG	TGTA	AAATTC	TCTAG	AAATTC	TATAT	GA	AA	1080
Db	1021	TGTC	CAAAAGCTG	GAAG	CAAA	CAAC	TCC	CAGTGGTG	TGTA	AAATTC	TCTAG	AAATTC	TATAT	GA	AA	1080
Qy	1081	AATG	ATGAATCTG	AGGTG	GA	AAATTTAT	TTCAT	GTGTAC	ATCCCTAT	CTCTG	TG	GA	AA	CAAG	GAGG	1140
Db	1081	AATG	ATGAATCTG	AGGTG	GA	AAATTTAT	TTCAT	GTGTAC	ATCCCTAT	CTCTG	TG	GA	AA	CAAG	GAGG	1140
Qy	1141	GCAG	ATTCANT	CCGTTAT	CTCTA	AAACAG	GTG	TACAGCA	AAATCCT	TAAAGT	CAC	TTTTA	AGATG			1200
Db	1141	GCAG	ATTCANT	CCGTTAT	CTCTA	AAACAG	GTG	TACAGCA	AAATCCT	TAAAGT	CAC	TTTTA	AGATG			1200
Qy	1201	TCAG	AAATAATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	1260
Db	1201	TCAG	AAATAATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	ATG	1260
Qy	1261	CAAC	CTTTT	TGAT	TTCTAT	TTG	AGGAGT	TGA	ATATAT	TG	CCAG	AGCTG	GA	CT	CGCT	1320
Db	1261	CAAC	CTTTT	TGAT	TTCTAT	TTG	AGGAGT	TGA	ATATAT	TG	CCAG	AGCTG	GA	CT	CGCT	1320
Qy	1321	GAGG	AAATAT	TGCT	TGGT	CCAT	CT	ACTAG	ATG	CGCT	CC	CAG	ACT	CG	CT	1380
Db	1321	GAGG	AAATAT	TGCT	TGGT	CCAT	CT	ACTAG	ATG	CGCT	CC	CAG	ACT	CG	CT	1380
Qy	1381	TTG	ATCT	CAC	CTG	AAATTTAT	TAT	CC	CAGT	AAG	ATG	ATG	TAT	TG	GA	1440
Db	1381	TTG	ATCT	CAC	CTG	AAATTTAT	TAT	CC	CAGT	AAG	ATG	ATG	TAT	TG	GA	1440
Qy	1441	ATTG	AGT	CAG	TG	CC	TG	AT	CTG	TG	AG	CG	CA	CT	AA	1500
Db	1441	ATTG	AGT	CAG	TG	CC	TG	AT	CTG	TG	AG	CG	CA	CT	AA	1500
Qy	1501	TG	GA	TA	AA	TAT	CC	AT	CA	CT	TT	TT	TT	TT	TT	1560
Db	1501	TG	GA	TA	AA	TAT	CC	AT	CA	CT	TT	TT	TT	TT	TT	1560
Qy	1561	TTT	AT	TTTT	TG	CC	T	GA	AT	CA	AA	AG	T	T	T	1620
Db	1561	TTT	AT	TTTT	TG	CC	T	GA	AT	CA	AA	AG	T	T	T	1620
Qy	1621	TTA	AG	GA	AA	CA	AA	TAT	AA	CA	GN	T	CA	GT	G	1680
Db	1621	TTA	AG	GA	AA	CA	AA	TAT	AA	CA	GN	T	CA	GT	G	1680
Qy	1681	TG	T	CT	TAT	CA	AG	AG	GAT	AG	CA	AA	T	CA	AG	1740
Db	1681	TG	T	CT	TAT	CA	AG	AG	GAT	AG	CA	AA	T	CA	AG	1740
Qy	1741	GG	AT	CT	TA	TAT	CC	AA	GT	TG	AT	CA	AG	GT	G	1800
Db	1741	GG	AT	CT	TA	TAT	CC	AA	GT	TG	AT	CA	AG	GT	G	1800
Qy	1801	T	CC	CT	TT	T	AG	CA	T	C	CT	AG	T	CA	G	1860
Db	1801	T	CC	CT	TT	T	AG	CA	T	C	CT	AG	T	CA	G	1860
Qy	1861	CT	GA	CT	G	CC	GT	CT	CA	CT	CT	T	G	CT	CA	1920
Db	1861	CT	GA	CT	G	CC	GT	CT	CA	CT	CT	T	G	CT	CA	1920
Qy	1921	AG	T	AG	CT	A	P	AG	TA	CC	CA	AG	AA	T	T	1980
Db	1921	AG	T	AG	CT	A	P	AG	TA	CC	CA	AG	AA	T	T	1980
Qy	1981	GA	AG	AT	G	A	C	CA	CT	T	G	CC	CA	CA	AG	2040
Db	1981	GA	AG	AT	G	A	C	CA	CT	T	G	CC	CA	CA	AG	2040
Qy	2041	C	CT	CT	CT	G	A	CT	A	P	AG	TA	CC	CA	AG	2100

QY	2101	TTGTAATGGGATGCTTACAGACCTCATGATCTACAGCCTGGAAGAAATATCCTACTGTG	2160
Db	2101	TTGTAATGGGATGCTTACAGACCTCATGATCTACAGCCTGGAAGAAATATCCTACTGTG	2160
QY	2161	CTGTTTCATATGTTGG--TCCTCAGGTGAGTGTGTAATATCGGTTTAAAGAGTCA	2218
Db	2161	CTGTTTCATATGTTGGTCTCCCTCAGGTGAGTGTGTAATATCGGTTTAAAGAGTCA	2220
QY	2219	AGTATTTCCGTTGAATACCTAGCCTCTCTAGGTTATGTTGTTAGTGATAGACAACA	2278
Db	2221	AGTATTTCCGTTGAATACCTAGCCTCTCTAGGTTATGTTGTTAGTGATAGACAACA	2280
QY	2279	GGGATCTCTGTCACCGAGGCTTAAATTTGAAGCGCCCTTTAAATATAAAATGGGTCAAA	2338
Db	2281	GGGATCTCTGTCACCGAGGCTTAAATTTGAAGCGCCCTTTAAATATAAAATGGGTCAAA	2340
QY	2339	TAGAAATGAGGATCAGGTGGAAGGACTCCAAATATCTAGCTCTCGATATGATTTCAATG	2398
Db	2341	TAGAAATGAGGATCAGGTGGAAGGACTCCAAATATCTAGCTCTCGATATGATTTCAATG	2400
QY	2399	ACTTAGATCGTGTGGGCATCCAGGCTGGTCTATGGAGGATACCTCTCCCTGATGGCAT	2458
Db	2401	ACTTAGATCGTGTGGGCATCCAGGCTGGTCTATGGAGGATACCTCTCCCTGATGGCAT	2460
QY	2459	TAATGCAGAGTCAGATATCTTCAGGCTTCTATGCTGGGGCCCGAGTCACTCTGTGGA	2518
Db	2461	TAATGCAGAGTCAGATATCTTCAGGCTTCTATGCTGGGGCCCGAGTCACTCTGTGGA	2520
QY	2519	TCCTTCATGATACAGGATACACGGAAGCTTATATGGTCACTGACAGAGTGAACAGG	2578
Db	2521	TCCTTCATGATACAGGATACACGGAAGCTTATATGGTCACTGACAGAGTGAACAGG	2580
QY	2579	GCTATTTACTTAGGATCTGTGGCCATGACGAGAAAGTCCGCTCTGACCAATCGTT	2638
Db	2581	GCTATTTACTTAGGATCTGTGGCCATGACGAGAAAGTCCGCTCTGACCAATCGTT	2640
QY	2639	TACTGCTCTTACATGTTTCTGATGAGAAATGTCATTTGACACATACCATATTTAC	2698
Db	2641	TACTGCTCTTACATGTTTCTGATGAGAAATGTCATTTGACACATACCATATTTAC	2700
QY	2699	TGAGTTTTTTAGTGAGGCTGGAAAGCCATATGATTTACAGATCTATCTCAGGAGAGAC	2758
Db	2701	TGAGTTTTTTAGTGAGGCTGGAAAGCCATATGATTTACAGATCTATCTCAGGAGAGAC	2760
QY	2759	ACAGCATAAGAGTTCTCGAATCGGGAGAACATTATGAACCTGCATCTTTTGCACTACCTTC	2818
Db	2761	ACAGCATAAGAGTTCTCGAATCGGGAGAACATTATGAACCTGCATCTTTTGCACTACCTTC	2820
QY	2819	AAGAAAACTTTGGATCAGTATTTGCTCTCTAAAAGTGATATAATTTTGACCTGTGTAGA	2878
Db	2821	AAGAAAACTTTGGATCAGTATTTGCTCTCTAAAAGTGATATAATTTTGACCTGTGTAGA	2880
QY	2879	ACTCTCTGTATACACTGGCTATTAAACCAATGAGGAGGTTTAAATCAACAGAAAAACACA	2938
Db	2881	ACTCTCTGTATACACTGGCTATTAAACCAATGAGGAGGTTTAAATCAACAGAAAAACACA	2940
QY	2939	GAATTGATCATCACATTTTGATACCTGCATGTAACATCTACTCTGAAATAAATGTGG	2998
Db	2941	GAATTGATCATCACATTTTGATACCTGCATGTAACATCTACTCTGAAATAAATGTGG	3000
QY	2999	TGCCATGAGGGCTCTAGGTTTGTGTTAGTAACTTAATACCTTAACCCCAATGCTCAA	3058
Db	3001	TGCCATGAGGGCTCTAGGTTTGTGTTAGTAACTTAATACCTTAACCCCAATGCTCAA	3060
QY	3059	AATCAAAATGATACATTTCTGAGAGACCGAGTAACCAATTAAGAAATTAATAAAAAAA	3118
Db	3061	AATCAAAATGATACATTTCTGAGAGACCGAGTAACCAATTAAGAAATTAATAAAAAAA	3120
QY	3119	AA 3120	
Db	3121	AA 3122	

AX354793

Sequence 1 from Patent WO0179473.

AX354793

AX354793.1 GI:18619526

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominiidae; Homo.

Meyers, R.A. and Williamson, M.

21953, a human prolyl oligopeptidase family member and uses thereof

Patent: WO 0179473-A 1 25-OCT-2001;

Millennium Pharmaceuticals, Inc. (US)

Location/Qualifiers

1. .3143

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

229. .2877

/note="unnamed protein product"

/codon_start=1

/protein_id="CAD22990.1"

/db_xref="GI:18619527"

/translation="MAAAETEQLGVEIFETADCEENISQDRPKLEPFYVERYSWQ

LKKLLADTRKYHYMAKAPHDHFMVKRNDPDGPHSDRIYIYLAAMSGENRENTLVSEI

PKTINRAAVLMSKWLPLDLFOATLDYGMTSREELLRKRKRITGVGIASDYHOSSG

TFLFOAGSGIYHVKGQPGQTOQLPRLNVLVETSCPNIRMDPKLCPADPDWAFIHSN

DIWISNIVTREERLTYVHNELANMEEDARSAGVATFLQEEFDRTYSGYMWCPKAEIT

PSGGKILRLVEENDESEVEIIVHTSPMLERRADSPRYPKTGTANPKVTFKMEITMI

DAGRIIDVIDKELIQPEILFEGVEIYARAGMTPGKVAWSILLDRSOTRQLQIVLIS

PELFTPVEDDDVNERQRLIESVPSVTPLIITYETTDIWINIHDIHFVFSQSEEELEF

IPASCKTPRHLVKITSLKESYKRSGLPAPSPDKPCPKIEETAITSGEWEVLGR

HGSIQNDVEVRLLVFEETKDSPLEHLYVVSYPGEVTRLDTRDGYSHSCCISQCHD

FTLSYKNQKPHCVSLYKLSPPEDPTCKEFTWATILDSAGLPDYTPPETLPSFES

TTGTLXLMYKPHDLQPKKYPTVLPIYGGPOVLVNNRFGKGVFRINTLASLGYV

VVVDNRGSGHRLKFGAPKYKMGQIEIDQVEGLQYLAASRFDLDRVGTGHWSY

GGYLSLMALMORSIDIPRATAGAPVTLMIFVDTGYTERYMGHPDQEQGYLGSVAMQ

AKKPSSEPNRLLLHGLFDENHVEHTSILLSFLVRAGKFPYDLQIYQERHSIRVPE

GEHYELHLLHVLQENLGSRIAALKVI"

Query Match

Best Local Similarity

Matches 3107; Conservative

99.5%; Score 3105.2; DB 6; Length 3143;

99.9%; Pred. No. 0;

0; Mismatches 3; Indels 0; Gaps 0;

QY

11

GCCTCGAGGCAAGCGCGCTGCTACTGCCCCCGCTGCTTCTTGTGCGCGTCCCGC

70

Db

26

GCCTCGAGGCAAGCGCGCTGCTACTGCCCCCGCTGCTTCTTGTGCGCGTCCCGC

85

QY

71

CTGGGTTGTACCGCGCGCGCGCGAGGAAAGCCACTGCAACAGGACCGAGTGGAGC

130

Db

86

CTGGGTTGTACCGCGCGCGCGCGAGGAAAGCCACTAACCAGGACCGAGTGGAGC

145

QY

131

GGCGCAGCATGAAGCGCGCGCGCGCTCCATAGCGCACGTCGGGACGTCGCGCGGG

190

Db

146

GGCGCAGCATGAAGCGCGCGCGCGCTCCATAGCGCACGTCGGGACGTCGCGCGGG

205

QY

191

GCCGCGGGGAAAGAAAATGCAACATGGCAGCAGCAATGGAACAGAACAGCTGGGTG

250

Db

206

GCCGCGGGGAAAGAAAATGCAACATGGCAGCAGCAATGGAACAGAACAGCTGGGTG

265

QY

251

AGATATTTCAAACTGCGGACTGTGAGGAGAAATATTGAATCACAGGATCGGCTAAATG

310

Db

266

AGATATTTCAAACTGCGGACTGTGAGGAGAAATATTGAATCACAGGATCGGCTAAATG

325

QY

311

AGCCTTTTATGTTGAGCGGTATTCCTCGAGTTCAGCTTAAAAAGCTGCTGCCGATACCA

370

Db

326

AGCCTTTTATGTTGAGCGGTATTCCTCGAGTTCAGCTTAAAAAGCTGCTGCCGATACCA

385

371	QY	GA	A	A	A	T	A	T	A	T	G	G	C	T	A	C	A	T	G	A	T	G	G	C	A	A	G	C	C	A	C	A	T	G	A	T	T	C	A	T	T	T	G	T	G	A	G	G	A	430
386	DB	GA	A	A	A	T	A	T	A	T	G	G	C	T	A	C	A	T	G	A	T	G	G	C	A	A	G	C	C	A	C	A	T	G	A	T	T	C	A	T	T	T	G	T	G	A	G	G	A	445
431	QY	A	T	G	A	T	C	C	A	G	A	T	G	G	A	C	T	C	A	T	C	A	G	A	G	A	T	C	T	A	T	T	A	C	C	T	G	C	C	A	T	G	T	G	T	G	A	G	A	450
446	DB	A	T	G	A	T	C	C	A	G	A	T	G	G	A	C	T	C	A	T	C	A	G	A	G	A	T	C	T	A	T	T	A	C	C	T	G	C	C	A	T	G	T	G	T	G	A	G	A	505
491	QY	G	A	G	A	A	A	T	A	C	A	C	T	G	T	T	T	T	A	T	T	C	T	G	A	A	T	C	C	C	A	A	A	A	A	C	T	A	T	C	A	A	T	A	G	A	G	A	550	
506	DB	G	A	G	A	A	A	T	A	C	A	C	T	G	T	T	T	A	T	T	C	T	G	A	A	T	C	C	C	A	A	A	A	A	A	C	T	A	T	C	A	A	T	A	G	A	G	A	565	
551	QY	T	G	C	T	C	T	T	G	A	G	C	C	T	T	T	T	T	T	T	T	G	S	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	610	
566	DB	T	G	C	T	C	T	T	G	A	G	C	C	T	T	T	T	T	T	T	T	T	G	S	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	625	
611	QY	C	T	G	A	G	A	G	A	A	C	T	A	T	T	A	G	A	G	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	670		
626	DB	C	T	G	A	G	A	G	A	A	C	T	A	T	T	A	G	A	G	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	685		
671	QY	A	C	G	A	T	T	A	C	C	A	G	A	G	A	G	A	G	A	C	A	T	T	C	T	T	T	C	A	G	C	G	T	A	G	T	A	G	A	T	T	A	T	A	C	A	G	730		
686	DB	A	C	G	A	T	T	A	C	C	A	G	A	G	A	G	A	C	A	T	T	C	T	T	T	C	A	G	C	G	T	A	G	T	A	G	A	T	T	A	T	A	C	A	G	745				
731	QY	T	A	A	A	G	A	G	A	G	G	C	C	A	A	G	A	T	T	A	G	C	A	C	A	C	A	C	C	T	T	T	A	A	G	C	C	A	A	T	C	A	G	A	A	790				
746	DB	T	A	A	A	G	A	G	A	G	G	C	C	A	A	G	A	T	T	A	G	C	A	C	A	C	C	T	T	T	A	A	G	C	C	A	A	T	C	A	G	A	A	805						
791	QY	C	T	A	G	T	T	G	T	C	C	A	A	C	A	T	A	C	G	A	T	G	A	T	G	C	C	C	C	T	G	A	T	C	A	C	A	G	A	C	T	G	A	T	G	850				
806	DB	C	T	A	G	T	T	G	T	C	C	A	A	C	A	T	A	C	G	A	T	G	A	T	G	C	C	C	C	T	G	A	T	C	A	C	A	G	A	C	T	G	A	T	G</					

Db	1466		TGCGTGAATCTGTGAGCCCACTAATATCTATGAGAAACAACAGACATCTGGATAATA	1525
Qy	1511		TCCATGACATCTTTTCATGTTTTTCCCAAAAGTCAAGAAGAAATGAGTTTATTTTGT	1570
Db	1526		TCCATGACATCTTTTCATGTTTTTCCCAAAAGTCAAGAAGAAATGAGTTTATTTTGT	1585
Qy	1571		CCTCTGAATGCAAAAACAGGTTTCGTCATTTATACAAAATTACATCTATTTTAAAGGAA	1630
Db	1586		CCTCTGAATGCAAAAACAGGTTTCGTCATTTATACAAAATTACATCTATTTTAAAGGAA	1645
Qy	1631		GCAATATATAACGATCCAGTGGTGGGCTGCTCTCCAAGTGAATTTCAAGTGTCTATCA	1690
Db	1646		GCAATATATAACGATCCAGTGGTGGGCTGCTCTCCAAGTGAATTTCAAGTGTCTATCA	1705
Qy	1691		AAGAGGAGATAGCAATATACAGTGTGCAATGCGAAGTTCTTGCCGGCATGGATCTAATA	1750
Db	1706		AAGAGGAGATAGCAATATACAGTGTGCAATGCGAAGTTCTTGCCGGCATGGATCTAATA	1765
Qy	1751		TCCAAGTTGATGAAGTCAGAAAGGCTGGTATATTTTGAAGGCCACCAAGACTCCCTTTAG	1810
Db	1766		TCCAAGTTGATGAAGTCAGAAAGGCTGGTATATTTTGAAGGCCACCAAGACTCCCTTTAG	1825
Qy	1811		AGCATCACCTGTACGTAGTCAGTTACGTAAATCCTGGAGAGGTGACAAGGCTGACTGACC	1870
Db	1826		AGCATCACCTGTACGTAGTCAGTTACGTAAATCCTGGAGAGGTGACAAGGCTGACTGACC	1885
Qy	1871		GTGGCTACTCACATTTCTTGTCATCAGTCAGACACTGTGACTCTTTTAAAGTAAAGTATA	1930
Db	1886		GTGGCTACTCACATTTCTTGTCATCAGTCAGACACTGTGACTCTTTTAAAGTAAAGTATA	1945
Qy	1931		GTAAACAGAAGAAATCCACACTGTGTGTCCTTTTACAAGCTATCAAGCTCTGGAAGTACC	1990
Db	1946		GTAAACAGAAGAAATCCACACTGTGTGTCCTTTTACAAGCTATCAAGCTCTGGAAGTACC	2005
Qy	1991		CAACTTTCGAAAACAAAGGAAATTTTGGGCCACCAATTTTGGATTACAGCAGGCTCTCTCGT	2050
Db	2006		CAACTTTCGAAAACAAAGGAAATTTTGGGCCACCAATTTTGGATTACAGCAGGCTCTCTCGT	2065
Qy	2051		ACTATACTCCTCCAGAAATTTTCTCTTTGAAAGTACTACTGGATTACATTTGATGGGA	2110
Db	2066		ACTATACTCCTCCAGAAATTTTCTCTTTGAAAGTACTACTGGATTACATTTGATGGGA	2125
Qy	2111		TGCTCTACAAGCTCATGATCTACAGGCTCGAAAAGAAATCCTACTGTGCTGTTCAAT	2170
Db	2126		TGCTCTACAAGCTCATGATCTACAGGCTCGAAAAGAAATCCTACTGTGCTGTTCAAT	2185
Qy	2171		ATGGTGGTCTCCTCAGGTGCAAGTGGTGAATAATCGGTTTAAAGAGTCAAGTATTTCCGCT	2230
Db	2186		ATGGTGGTCTCCTCAGGTGCAAGTGGTGAATAATCGGTTTAAAGAGTCAAGTATTTCCGCT	2245
Qy	2231		TGAATACCTTAGGCTCTCTAGGTTATGTGGTTGTAGTGATAGACACAGGGGATCCTGTC	2290
Db	2246		TGAATACCTTAGGCTCTCTAGGTTATGTGGTTGTAGTGATAGACACAGGGGATCCTGTC	2305
Qy	2291		ACCGAGGGCTTAAATTTGAAAGGCGCTTTAAATATAAAATGGGCTCAAAATAGAAAATTGAGC	2350
Db	2306		ACCGAGGGCTTAAATTTGAAAGGCGCTTTAAATATAAAATGGGCTCAAAATAGAAAATTGAGC	2365
Qy	2351		ATCAGGTGGAAGGACTCCAAATATCTAGCTTCTCGATATGATTTTCAATTGACTTAGATCCGTG	2410
Db	2366		ATCAGGTGGAAGGACTCCAAATATCTAGCTTCTCGATATGATTTTCAATTGACTTAGATCCGTG	2425
Qy	2411		TGGGCATCCACGGCTGCTCTATGGAGGATACCTCTCCCTGATGGCAATTAATGCGAGAGT	2470
Db	2426		TGGGCATCCACGGCTGCTCTATGGAGGATACCTCTCCCTGATGGCAATTAATGCGAGAGT	2485
Qy	2471		CAGATATCTTCAGGGTTGCTATTGCTGGGGGCCAGTCACTCTGTGGATCTTCTATGATA	2530
Db	2486		CAGATATCTTCAGGGTTGCTATTGCTGGGGGCCAGTCACTCTGTGGATCTTCTATGATA	2545
Qy	2531		CAGGATACCGGAACGTTATATGGGTCACCTCTGACCAAGATGAACAGGGCTATTACTTAG	2590

Db	2546	CAGGATACAGGAACGTTATATGGGTACCCCTGACACAGAAATGAA CAGGGCTATTACTTTAG	2605
Qy	2591	GATCTGTGGCCATCAAGACGAGAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTAC	2650
Db	2606	GATCTGTGGCCATCAAGACGAGAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTAC	2665
Qy	2651	ATGGTTTCTCTGGATGAGAATGTCCATTTTGCACATACCAAGTATATTAATCTGAGTTTTTTAG	2710
Db	2666	ATGGTTTCTCTGGATGAGAATGTCCATTTTGCACATACCAAGTATATTAATCTGAGTTTTTTAG	2725
Qy	2711	TGAGGGCTGGAAAGCCATATGATTTACAGATCTATCTCTCAGAGAGACACAGCATAAGAG	2770
Db	2726	TGAGGGCTGGAAAGCCATATGATTTACAGATCTATCTCTCAGAGAGACACAGCATAAGAG	2785
Qy	2771	TTCCCTGAATCGGAGACATATTAAGCTGATTAATTTTGGCACTCTTTGCACTACCTTCAAGAAACCTTG	2830
Db	2786	TTCCCTGAATCGGAGACATATTAAGCTGATTAATTTTGGCACTCTTTGCACTACCTTCAAGAAACCTTG	2845
Qy	2831	GATCAGTATTGCTGCTCTTAAGAGTGATTAATTTTGGCACTCTTTGCACTACCTTCAAGAAACCTTG	2890
Db	2846	GATCAGTATTGCTGCTCTTAAGAGTGATTAATTTTGGCACTCTTTGCACTACCTTCAAGAAACCTTG	2905
Qy	2891	ACACTGGCTATTAAACCAATGAGGAGTTTAACTCAACAGAAAAACAGAAATGATCATC	2950
Db	2906	ACACTGGCTATTAAACCAATGAGGAGTTTAACTCAACAGAAAAACAGAAATGATCATC	2965
Qy	2951	ACATTTTGATACCTGCCATGTAAACATCTCTCTGAAAAATAATGTGGTCCATCGAGG	3010
Db	2966	ACATTTTGATACCTGCCATGTAAACATCTCTCTGAAAAATAATGTGGTCCATCGAGG	3025
Qy	3011	GTCTACGGTTTGTGTAGTAACTTAATACCTTAAACCCCATCTCAAAATCAAAATGATA	3070
Db	3026	GTCTACGGTTTGTGTAGTAACTTAATACCTTAAACCCCATCTCAAAATCAAAATGATA	3085
Qy	3071	CATATTCTCTGAGAGACCCAGCAATACCAATAGAAATTAATAAAAAA	3120
Db	3086	CATATTCTCTGAGAGACCCAGCAATACCAATAGAAATTAATAAAAAA	3135
RESULT 6			
AX342633			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
Homo sapiens (human)			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;			
Hominidae; Homo.			
1			
REFERENCE			
AUTHORS			
Yue, H., Elliott, V.S., Gandhi, A.R., Lal, P., Au-Young, J.,			
Tribouley, C.M., Deleage, A.M., Baughn, M.R., Nguyen, D.B., Lee, E.A.,			
Hafalia, A., Khan, F.A., Wallia, N.K., Yao, M.G., Lu, D.A., Patterson, C.,			
Tang, Y.T., Walsh, R.T., Azimzai, Y., Ramkumar, J., Xu, Y. and Reddy, R.			
Patent: WO 0198468-A 30 27-DEC-2001;			
Incyte Genomics, Inc. (US)			
FEATURES			
Location/Qualifiers			
1..3106			
/organism="Homo sapiens"			
/mol_type="unassigned DNA"			
/db_xref="taxon:9606"			
/note="Incyte ID No: 7160544CB1"			
ORIGIN			
Query Match 99.4%; Score 3101.8; DB 6; Length 3106;			
Best Local Similarity 99.9%; Pred. No. 0;			
Matches 3103; Conservative 0; Mismatches 2; Indels 0; Gaps 0;			
Qy	13	CTCCGAGGCGCAAGGCGGCTACTGCGCGCGCTGCTTCTTAGTGGCGGCTTCGCGGCT	72
Db	2	CTCCGAGGCGCAAGGCGGCTACTGCGCGCGCTGCTTCTTAGTGGCGGCTTCGCGGCT	61

1153 CGTTATCTCTAAACAGGTACAGCAAAATCCTAAAGTCACTTTTAAAGATGTGAGAAATATG 1212
1142 CGTTATCTCTAAACAGGTACAGCAAAATCCTAAAGTCACTTTTAAAGATGTGAGAAATATG 1201
1213 ATTGATGCTGAAGGAAGATCATAGATGTCTATAGATGAAGAACTAAATCAACCTTTGAG 1272
1202 ATTGATGCTGAAGGAAGATCATAGATGTCTATAGATGAAGAACTAAATCAACCTTTGAG 1261
1273 ATTCTATTGGAAGGAGTTGAATATATTGCCAGAGCTGGATGGAATCTCTGAGGGAATAAT 1332
1262 ATTCTATTGGAAGGAGTTGAATATATTGCCAGAGCTGGATGGAATCTCTGAGGGAATAAT 1321
1333 GCTGGTGCATCTCTACTAGATGCTCCAGAGCTGCCCTACAGATAGTGTGATCTCACCT 1392
1322 GCTGGTGCATCTCTACTAGATGCTCCAGAGCTGCCCTACAGATAGTGTGATCTCACCT 1381
1393 GAATATTATTATCCAGTAGAAGATGATGTTATGGAAGGAGAGACTCAATTGAGTCAGTG 1452
1382 GAATATTATTATCCAGTAGAAGATGATGTTATGGAAGGAGAGACTCAATTGAGTCAGTG 1441
1453 CCTGATTTCTGTGACGCCACTAATTATCTATGAAGAAACAAACAGACATCTGGATAATATC 1512
1442 CCTGATTTCTGTGACGCCACTAATTATCTATGAAGAAACAAACAGACATCTGGATAATATC 1501
1513 CATGACATCTTTTATGTTTTTCCCAAGATCGAAGAGGAATGAGTTTATTTTGGC 1572
1502 CATGACATCTTTTATGTTTTTCCCAAGATCACGAGAGGAATGAGTTTATTTTGGC 1561
1573 TCTGAATGCAAAAGAGTTTCGTCATTTATACAAATTAATCAATTTTAAAGGAAGC 1632
1562 TCTGAATGCAAAAGAGTTTCGTCATTTATACAAATTAATCAATTTTAAAGGAAGC 1621
1633 AAATATAAACGATCAGTGGTGGGCTGCTGCTCCAAAGTGAATTTCAAGTGTCTCTACAA 1692
1622 AAATATAACGATCAGTGGTGGGCTGCTGCTCCAAAGTGAATTTCAAGTGTCTCTACAA 1681
1693 GAGGAGATAGCAATTAACAGTGGTGAATGGGAAGTTCTTGCCGGGATGATCTAATATC 1752
1682 GAGGAGATAGCAATTAACAGTGGTGAATGGGAAGTTCTTGCCGGGATGATCTAATATC 1741
1753 CAAGTTGATGAAGTCAAGAGGTGGTATATTTTGAAGGCCAACAGCTCCCTTTAGAG 1812
1742 CAAGTTGATGAAGTCAAGAGGTGGTATATTTTGAAGGCCAACAGCTCCCTTTAGAG 1801
1813 CATCACTGTACGTAGTCAAGTACGTAAATCTCGAGAGGTGCAAGGCTGACGACCGT 1872
1802 CATCACTGTACGTAGTCAAGTACGTAAATCTCGAGAGGTGCAAGGCTGACGACCGT 1861
1873 GGCTACTCACATTTCTGTGATCAGTCAGCACTGTGACTTCTTTTAAAGTAAAGTATAGT 1932
1862 GGCTACTCACATTTCTGTGATCAGTCAGCACTGTGACTTCTTTTAAAGTAAAGTATAGT 1921
1933 AACGAGAAGATCCACATGTGTGTCCTTTTAAAGCTATCAAGCTCAAGTCTGAAGATGACCCA 1992
1922 AACGAGAAGATCCACATGTGTGTCCTTTTAAAGCTATCAAGCTCAAGTCTGAAGATGACCCA 1981
1993 ACTTGCAGAAACAAAGGAATTTTGGCCACCAATTTTGGATTCAGCAGCTCTCTTCTGAC 2052
1982 ACTTGCAGAAACAAAGGAATTTTGGCCACCAATTTTGGATTCAGCAGCTCTCTTCTGAC 2041
2053 TATACTCTCTCAGAAATTTTCTCTTTTGAAGTACTACTGGATTTTCAATTTGATGGATG 2112
2042 TATACTCTCTCAGAAATTTTCTCTTTTGAAGTACTACTGGATTTTCAATTTGATGGATG 2101
2113 CTCTACAGCCTCATGATCTACAGCTCGGAAAGAAATATCTTACTGTGCTGTTTCAATAT 2172
2102 CTCTACAGCCTCATGATCTACAGCTCGGAAAGAAATATCTTACTGTGCTGTTTCAATAT 2161
2173 GGTGGTCTCAGGTCAGTTGGTGAATATCGTTTAAAGAGTCAAGTATTTCCGCTTG 2232
2162 GGTGGTCTCAGGTCAGTTGGTGAATATCGTTTAAAGAGTCAAGTATTTCCGCTTG 2221
2233 AATACCTAGCCTCTCTAGGTTATGTTGGTTGATGATAGACAAAGGGATCCTGTCAC 2292

Db 2222 AATACCTAGCCTCTCTAGGTTATGTTGGTTAGTATAGACAAAGGGATCCTGTCAC 2281
Qy 2293 CGAGGCTTTAAATTTGAAGGCGCTTTAAATATATAAATGGGTCAAAATAGAAATGACAT 2352
Db 2282 CGAGGCTTTAAATTTGAAGGCGCTTTAAATATATAAATGGGTCAAAATAGAAATGACAT 2341
Qy 2353 CAGGTGGAAGGATCCAATATCTAGCTTCTCGATATGATTTTCATTTGATCTAGATCGGTG 2412
Db 2342 CAGGTGGAAGGATCCAATATCTAGCTTCTCGATATGATTTTCATTTGATCTAGATCGGTG 2401
Qy 2413 GGCATCCAGGCTGGTCTTATGGAGGATACCTCTCCCTGATGGCATTAATGACAGAGTCA 2472
Db 2402 GGCATCCAGGCTGGTCTTATGGAGGATACCTCTCCCTGATGGCATTAATGACAGAGTCA 2461
Qy 2473 GATATCTTCAGGCTTGTCTATTTGCGGGCCCCAGTCACTCTGTGGATCTTCTATGATACA 2532
Db 2462 GATATCTTCAGGCTTGTCTATTTGCGGGCCCCAGTCACTCTGTGGATCTTCTATGATACA 2521
Qy 2533 GATACACGGAAGCTTATATGGGTCACTCTGACCAAGATGAACAGGGCTATTACTTAGGA 2592
Db 2522 GATACACGGAAGCTTATATGGGTCACTCTGACCAAGATGAACAGGGCTATTACTTAGGA 2581
Qy 2593 TCTGTGCCATGCAAGCAGAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACAT 2652
Db 2582 TCTGTGCCATGCAAGCAGAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACAT 2641
Qy 2653 GGTTCCTCGATGAGAAATGTCATTTTGCACATACAGTATATTTACTGAGTTTTTATAGT 2712
Db 2642 GGTTCCTCGATGAGAAATGTCATTTTGCACATACAGTATATTTACTGAGTTTTTATAGT 2701
Qy 2713 AGGGCTGGAAGCATAATGATTTACAGATCTATCTCAGGAGAGACACAGCATTAAGATT 2772
Db 2702 AGGGCTGGAAGCATAATGATTTACAGATCTATCTCAGGAGAGACACAGCATTAAGATT 2761
Qy 2773 CCTGAATCGGAGAACATTTAGAACTGTCATCTTTTGGACTACCTTTCAAGAAACCTTTGGA 2832
Db 2762 CCTGAATCGGAGAACATTTAGAACTGTCATCTTTTGGACTACCTTTCAAGAAACCTTTGGA 2821
Qy 2833 TCACGTATTGCTGCTCTAAAGTGATATAATTTTGACCTGTGTAGAACTCTCTGCTATAC 2892
Db 2822 TCACGTATTGCTGCTCTAAAGTGATATAATTTTGACCTGTGTAGAACTCTCTGCTATAC 2881
Qy 2893 ACTGGCTATTAAACCAATGAGGAGTTTAAATCAACAGAAACACAGAAATGATCATCAC 2952
Db 2882 ACTGGCTATTAAACCAATGAGGAGTTTAAATCAACAGAAACACAGAAATGATCATCAC 2941
Qy 2953 ATTTTGATACCTGCGCATGTAAACATCTACTCTGAAATTAATGTGGTCCCATGAGGGT 3012
Db 2942 ATTTTGATACCTGCGCATGTAAACATCTACTCTGAAATTAATGTGGTCCCATGAGGGT 3001
Qy 3013 CTACGTTTGTGTAGTAAATCTAAATCACTTAAACCCACATGCTCAAAATCAAAATGATACA 3072
Db 3002 CTACGTTTGTGTAGTAAATCTAAATCACTTAAACCCACATGCTCAAAATCAAAATGATACA 3061
Qy 3073 TATTCTGAGAGACCCAGCAATACCATAAGAAATTAATAAAAAA 3117
Db 3062 TATTCTGAGAGACCCAGCAATACCATAAGAAATTAATAAAAAA 3106

RESULT 7
BC040203

LOCUS Homo sapiens dipeptidylpeptidase 8, mRNA (cdna clone MGC:42570
DEFINITION IMAGE:4824813), complete cds.

ACCESSION BC040203

VERSION BC040203.1 GI:26007916

KEYWORDS MGC.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.

REFERENCE AUTHORS	1 (bases 1 to 4535) Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altshul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Heieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.P., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,W.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butlerfield,V., Skrzewinski,M.I., Skalska,U., Smallos,D.E., Schnerich,A., Schein,J.E., Jones,S.J. and Marra,M.A.
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
PUBMED	12477932
REFERENCE AUTHORS	2 (bases 1 to 4535) Strausberg,R. Direct Submission
TITLE	Submitted (27-NOV-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
JOURNAL	NIH-MGC Project URL: http://mgc.nci.nih.gov Contact: MGC help desk Email: cgapbs-remail.nih.gov Tissue procurement: Miklos Palkovits, M.D., Ph.D. cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki Toshiyuki and Piero Carninci (RIKEN) cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center Center code: BCM-HGSC Web site: http://www.hgsc.bcm.tmc.edu/cdna/ Contact: amg@bcm.tmc.edu Gunaratne P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H., Kowis, C.R., Sneed, A.J., Martin, R.G., Wuzny, D.M., Nanavati, A.N., Gibbs, R.A.
REMARK COMMENT	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Series: IRAK Place: 71 Row: K Column: 23. Location/Qualifiers 1. .4535 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="MGC:42570 IMAGE:4824813" /tissue_type="Testis" /clone_lib="NIH MGC_97" /lab_host="DH10B" /note="Vector: pBluescript" 1. .4535 /gene="Dpp8" /note="synonyms: FLJ20283, FLJ14920, MGC26191" /db_xref="GeneID:54878" /db_xref="MIM:606819" 1629. .4277 /gene="Dpp8" /codon_start=1 /product="Dpp8 protein" /protein_id="AAH40203.1" /db_xref="GI:26007917" /db_xref="GeneID:54878" /db_xref="MIM:606819" /translation="MAAAETEQLGVEIFETADCEENIESQDRPKLEPFYVERYSWSQ
FEATURES source	LKKLLADTRKVGVMKAPHDFMVKRNDPDGPHSDRIYLYLAMSENRENTLFYSEI PKTINRAALMLSKPLLDLPQATLOYGMVSEERLLRKRKLTGTVGLASYDYHOGSS TFPLQAQSGLYHYKDGPOGFTOOLRPNLIVETSCPNRMKLPKPADPDWIAFIHGN DWLSNIVTREERLFTVNNELANMEDSAGVATFVLOESPDYSGYMWCPKAEIT PSGKILRLIENEDESEVEIIVHTSPMLTETRADSPRYKGTGTANPKVTFKMSIMI DAEGRILIDIKELIQFELFEGFIFIAAGWTGPGKYAMSLIDRSQTRLOIQLVIS PELFIPEVDDEMRQRUESVPSPVFIPIIIEITIIWINIHDIHFVHQSEHEERLEP IFASECKTGFRHLYKTSILUKSKYKSSGGLPAPSDPKCPKBEITATISGSEWELER HGSNIQDVEURLVYFEGTDSPLEHHLVYVSVNPGVETRLTRDGRKTFRLNTLASGLY FFIKYQNKPNHPCVLSKJSSPEDDPTCKTEFWALILDSAGKFKFRLNTLASGLY TVGTGLXLMYKHLQPLQKXYPVFIYGGPQVQVNNRKGKVFYKFLRLNTLASGLY VVVDNRGSCHRGLKFGKAPKMGQIEIDQVEGLOYLASRYDFIDDLQVYLGWSVY CGYLSLMALMQRSDFRVAIAGAVTLMIPFYDTGYTYRMYGHPDQNEQGYLGVSVMQ AEKFPSEFNLHLLHGFEDBNVFAHTSILSLFLVRAGKPYDLOIYQPERHSIRVPS GEHYELHLLHYLQENLGSRTAALKVI" 3207. .4196 /gene="Dpp8" /note="DAP2; Region: Dipeptidyl aminopeptidases/acylaminoacyl-peptidases (Amino acid transport and metabolism)" /db_xref="CDD:COG1506"
misc_feature	
ORIGIN	Query Match 93.4%; Score 2915.2; DB 8; Length 4535; Best Local Similarity 99.9%; Pred. No. 0; Matches 2917; Conservative 0; Mismatches 3; Indels 0; Gaps 0; Qy 201 AGGAAATGCAACATGGCAGCAGCAATGGAAACAGACAGCTGGTGTTCGAGATATTTGA 260 Db 1616 AGGAAATGCAACATGGCAGCAGCAATGGAAACAGACAGCTGGTGTTCGAGATATTTGA 1675 Qy 261 AACTGGGACTGTGAGGAGATATGATACAGATCGGCCTAAATTCGAGCCTTTTGA 320 Db 1676 AACTGGGACTGTGAGGAGATATGATACAGATCGGCCTAAATTCGAGCCTTTTGA 1735 Qy 321 TGTGAGCGGTATTCCTGGAGTACGCTTAAAGCTGTTCCGATACAGAAATATCA 380 Db 1736 TGTGAGCGGTATTCCTGGAGTACGCTTAAAGCTGTTCCGATACAGAAATATCA 1795 Qy 381 TGCTACATGATGGCTAAGGGACCAACATGATTTTCAATGTTGTGAAGAGGAATGATCCAGA 440 Db 1796 TGCTACATGATGGCTAAGGGACCAACATGATTTTCAATGTTGTGAAGAGGAATGATCCAGA 1855 Qy 441 TGACCTCATTCAGACAGAGATCTTATTACCTTGGCATGTCTGGTGAAGAACAGAAATATC 500 Db 1856 TGACCTCATTCAGACAGAGATCTTATTACCTTGGCATGTCTGGTGAAGAACAGAAATATC 1915 Qy 501 ACTGTTTATTCGAAATTCCTGAAATTCCTGAGCAACACTCAATAGAGCAGCAGTCTTAATGCTCTT 560 Db 1916 ACTGTTTATTCGAAATTCCTGAAATTCCTGAGCAACACTCAATAGAGCAGCAGTCTTAATGCTCTT 1975 Qy 561 GAAGCCTCTTTTGGATCTTTTTCAGGCAACACTGAGCTATGGAATGATTTCTCGAGAAGA 620 Db 1976 GAAGCCTCTTTTGGATCTTTTTCAGGCAACACTGAGCTATGGAATGATTTCTCGAGAAGA 2035 Qy 621 AGAATCTTAAAG 680 Db 2036 AGAATCTTAAAG 2095 Qy 681 CCAAGGAAGTGAACATTTCTGTTTCAAGCGGTAGTGAATTTATCAGCTTAAAGATGG 740 Db 2096 CCAAGGAAGTGAACATTTCTGTTTCAAGCGGTAGTGAATTTATCAGCTTAAAGATGG 2155 Qy 741 AGGGCCACCAAGGATTTACGCAACACTTTAAGGCCCAATCTAGTGGAACTAGTCTGCC 800 Db 2156 AGGGCCACCAAGGATTTACGCAACACTTTAAGGCCCAATCTAGTGGAACTAGTCTGCC 2215 Qy 801 CAACATACCGATGGATTCAGAAATTTATGCCCGCTGATCCAGCTGGATTTGTTTTATACA 860 Db 2216 CAACATACCGATGGATTCAGAAATTTATGCCCGCTGATCCAGCTGGATTTGTTTTATACA 2275 Qy 861 TAGCAACGATATTTGGATATCTAACATCGTAACCAAGAGAAAGGAGGACTCACTTATGT 920

Db 2276 TGGCAACGATATTTGGATATCTAAATCGTAAACAGAGAAAGGAGACTCACTATGT 2335
Qy 921 GCACATAGCTAGCCAAATCGAAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTGT 980
Db 2336 GCACATAGCTAGCCAAATCGAAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTGT 2395
Qy 981 TCTCCAGAGAAATTTGATAGATATCTGGCTATTTGGTGTCTGGTGTCCAAAAGCTGAAACAAC 1040
Db 2396 TCTCCAGAGAAATTTGATAGATATCTGGCTATTTGGTGTCTGGTGTCCAAAAGCTGAAACAAC 2455
Qy 1041 TCCAGTGTGGTGAATAATCTTAGAAATCTATATGAAGAAATGATGAATCTGAGGTGA 1100
Db 2456 TCCAGTGTGGTGAATAATCTTAGAAATCTATATGAAGAAATGATGAATCTGAGGTGA 2515
Qy 1101 AATTATTCATGTTCATCCCTATATTTGGAAACAAGGAGGCGAGATTCATTCGGTATCC 1160
Db 2516 AATTATTCATGTTCATCCCTATATTTGGAAACAAGGAGGCGAGATTCATTCGGTATCC 2575
Qy 1161 TAAAAACAGGTACAGCAAAATCTTAAAGTCACTTTTAAAGATGTCAAGAAATATGATGTGC 1220
Db 2576 TAAAAACAGGTACAGCAAAATCTTAAAGTCACTTTTAAAGATGTCAAGAAATATGATGTGC 2635
Qy 1221 TGAAGGAGGATCATAGATGTCTAGATATAGGAATTAAGTCACTTTTGAAGATTCATTT 1280
Db 2636 TGAAGGAGGATCATAGATGTCTAGATATAGGAATTAAGTCACTTTTGAAGATTCATTT 2695
Qy 1281 TGAAGGAGGATTCATATATTTGACAGCTGGAATGACCTCTGAGGAAATATGCTTGGTC 1340
Db 2696 TGAAGGAGGATTCATATATTTGACAGCTGGAATGACCTCTGAGGAAATATGCTTGGTC 2755
Qy 1341 CATCTCTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACTCGAATTAAT 1400
Db 2756 CATCTCTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACTCGAATTAAT 2815
Qy 1401 TATCCAGTAGAAGATGATTTTGAAGAGCAGAGACTCAATGAGTCAGTCTGATTC 1460
Db 2816 TATCCAGTAGAAGATGATTTTGAAGAGCAGAGACTCAATGAGTCAGTCTGATTC 2875
Qy 1461 TGTGAGCCACTAATTTCTATGAAGAAACAACAGACATCTGGATAAATATCATGACAT 1520
Db 2876 TGTGAGCCACTAATTTCTATGAAGAAACAACAGACATCTGGATAAATATCATGACAT 2935
Qy 1521 CTTTCATGTTTTCCCAAGTCACGAAGAGAAATTTAGTTTTATTTTTGCCTCTGAATG 1580
Db 2936 CTTTCATGTTTTCCCAAGTCACGAAGAGAAATTTAGTTTTATTTTTGCCTCTGAATG 2995
Qy 1581 CAAAAACAGGTTTCCGTCAATTTATACAAAATTTACATCTATTTTAAAGAAAGCAATATAA 1640
Db 2996 CAAAAACAGGTTTCCGTCAATTTATACAAAATTTACATCTATTTTAAAGAAAGCAATATAA 3055
Qy 1641 ACGATCCAGTGTGGGTGCGCTGCTCCAAGTGATTTCAAGTGTCTATCAAGAGGAGAT 1700
Db 3056 ACGATCCAGTGTGGGTGCGCTGCTCCAAGTGATTTCAAGTGTCTATCAAGAGGAGAT 3115
Qy 1701 AGCAATTAACAGTGTGATGGGAATTTCTGGCCGCGATGGATCTAATATCCAAGTTGA 1760
Db 3116 AGCAATTAACAGTGTGATGGGAATTTCTGGCCGCGATGGATCTAATATCCAAGTTGA 3175
Qy 1761 TGAAGTCAGAAGGCTGGTATATTTTGAAGGCAACAAAGACTCCCTTTTAGAGCATCACCT 1820
Db 3176 TGAAGTCAGAAGGCTGGTATATTTTGAAGGCAACAAAGACTCCCTTTTAGAGCATCACCT 3235
Qy 1821 GTACGTAGTCAGTTACGTTAAATCTTGGAGAGGTGACAAGGCTGACTGACCGTGGCTACT 1880
Db 3236 GTACGTAGTCAGTTACGTTAAATCTTGGAGAGGTGACAAGGCTGACTGACCGTGGCTACT 3295
Qy 1881 ACATTCCTGCTGCATCAGTCAGCACTGTGACTCTTTTATAAGTAAGTATAGTAAACAGAA 1940
Db 3296 ACATTCCTGCTGCATCAGTCAGCACTGTGACTCTTTTATAAGTAAGTATAGTAAACAGAA 3355
Qy 1941 GAATCCACACTGTGTGCTCCTTTTACAAGCTATCAAGTCTTCAAGTCTGAGAGTACCAATCTTGCAA 2000
Db 3356 GAATCCACACTGTGTGCTCCTTTTACAAGCTATCAAGTCTTCAAGTCTGAGAGTACCAATCTTGCAA 3415

Qy 2001 AACAAAGGAATTTTGGGCCACCAATTTTGGATTTAGCAGGCTCCTCTTCTGACTATACTCC 2060
Db 3416 AACAAAGGAATTTTGGGCCACCAATTTTGGATTTAGCAGGCTCCTCTTCTGACTATACTCC 3475
Qy 2061 TCCAGAAAATTTTCTCTTTTGAAGTACTACTGGATTTACATTTGTATGGGATGCTCTACAA 2120
Db 3476 TCCAGAAAATTTTCTCTTTTGAAGTACTACTGGATTTACATTTGTATGGGATGCTCTACAA 3535
Qy 2121 GCCTCATGATCTACAGCCTGGAAAGAAATATCTACTGTCTGTGTTCATATATGTGTGTC 2180
Db 3536 GCCTCATGATCTACAGCCTGGAAAGAAATATCTACTGTCTGTGTTCATATATGTGTGTC 3595
Qy 2181 TCAGGTGCAATTTGGTGAATTAATTCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCT 2240
Db 3596 TCAGGTGCAATTTGGTGAATTAATTCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCT 3655
Qy 2241 AGCCTCTCTAGGTTATGTGTGTGTAGTATGACAAACAGGGGATCTGTCTCACCCAGAGCT 2300
Db 3656 AGCCTCTCTAGGTTATGTGTGTGTAGTATGACAAACAGGGGATCTGTCTCACCCAGAGCT 3715
Qy 2301 TAAATTTGAAGCGCCTTTTAAATATATAAATGGGTCAAAATAGAAAATTCACGATCAGGTGA 2360
Db 3716 TAAATTTGAAGCGCCTTTTAAATATATAAATGGGTCAAAATAGAAAATTCACGATCAGGTGA 3775
Qy 2361 AGGACTCCAATATCTAGCTTCTCGATATGATTTCAATTCAGTCTAGATCGTGTGGGATCCA 2420
Db 3776 AGGACTCCAATATCTAGCTTCTCGATATGATTTCAATTCAGTCTAGATCGTGTGGGATCCA 3835
Qy 2421 CGGCTGGTCTATGGAGGATACCTCTCCCTGATGTCATTAATGAGGTCAGAGTCAAGATCTTT 2480
Db 3836 CGGCTGGTCTATGGAGGATACCTCTCCCTGATGTCATTAATGAGGTCAGAGTCAAGATCTTT 3895
Qy 2481 CAGGTTGCTATTTGCTGGGCGCCAGTCACTCTGTGGATCTTCTATGATACAGATACAC 2540
Db 3896 CAGGTTGCTATTTGCTGGGCGCCAGTCACTCTGTGGATCTTCTATGATACAGATACAC 3955
Qy 2541 GGAACGTTATATGGGTCACTTGACAGAAAGCAAGGCTATTTACTTGAATCTGTGGC 2600
Db 3956 GGAACGTTATATGGGTCACTTGACAGAAAGCAAGGCTATTTACTTGAATCTGTGGC 4015
Qy 2601 CATGCAAGCAGAAAGTTCCCTCTGAAACCAATCTTCTGCTCTTACATGCTTTCCCT 2660
Db 4016 CATGCAAGCAGAAAGTTCCCTCTGAAACCAATCTTCTGCTCTTACATGCTTTCCCT 4075
Qy 2661 GGATGAGAAATGTCATTTTGGACATACCAAGTATATTTACTGAGTTTTTTTGTGAGGCTGG 2720
Db 4076 GGATGAGAAATGTCATTTTGGACATACCAAGTATATTTACTGAGTTTTTTTGTGAGGCTGG 4135
Qy 2721 AAAGCCATATGATTTTACAGATCTATCTCTCAGGAGAGACACAGCATAGAGTTCTGAAATC 2780
Db 4136 AAAGCCATATGATTTTACAGATCTATCTCTCAGGAGAGACACAGCATAGAGTTCTGAAATC 4195
Qy 2781 GGAGAAACATTTATGAACTCTGATCTTTTGGATCTACCTTCAAGAAACCTTGGATCAGGTAT 2840
Db 4196 GGAGAAACATTTATGAACTCTGATCTTTTGGATCTACCTTCAAGAAACCTTGGATCAGGTAT 4255
Qy 2841 TGCTGCTCTAAAGTATATATTTTGAACCTGTGTAGAACTCTCTGTGTATACATGCTGCTA 2900
Db 4256 TGCTGCTCTAAAGTATATATTTTGAACCTGTGTAGAACTCTCTGTGTATACATGCTGCTA 4315
Qy 2901 TTTTAAACCAATGAGGAGTTTAAATCAACAGAAAAACAGAAATTTGATCATCATTTTGTAT 2960
Db 4316 TTTTAAACCAATGAGGAGTTTAAATCAACAGAAAAACAGAAATTTGATCATCATTTTGTAT 4375
Qy 2961 ACCTGCCATGTAAATCTACTCTGAAATAAATGTGTGCTGATCAGAGGCTCTACGGTT 3020
Db 4376 ACCTGCCATGTAAATCTACTCTGAAATAAATGTGTGCTGATCAGAGGCTCTACGGTT 4435
Qy 3021 TGTGTGTATATCTTAATACCTTAAACCCACATGCTCAAAATCAAAATGATATATATCTCTG 3080
Db 4436 TGTGTGTATATCTTAATACCTTAAACCCACATGCTCAAAATCAAAATGATATATATCTCTG 4495

Qy	3081	AGAGCCAGCAACATACCAATAGATTTACTAATAAATAAATAA 3120	Db	601	CGAATGTATCTCGAGAGAGAACTATTAAAGAGAAAGAAACGCAATGGAAACAGTCGGA 660
Db	4496	AGAGCCAGCAACATACCAATAGATTTACTAATAAATAAATAA 4535		661	ATTGCTTCTTACGATTATCAACAAGGAAGTGGAAACATTTCTGTTTCAAGCCGTTAGTGA 720
				661	ATTGCTTCTTACGATTATCAACAAGGAAGTGGAAACATTTCTGTTTCAAGCCGTTAGTGA 720
RESULT 8				721	ATTATATCGTAAAGATGGAGGGCCCAAGGATTTACGCAACAACCTTTAAAGCCCAAT 780
LOCUS	AR631288	4685 bp DNA linear PAT 14-FEB-2005		721	ATTATATCGTAAAGATGGAGGGCCCAAGGATTTACGCAACAACCTTTAAAGCCCAAT 780
DEFINITION	Sequence 22 from patent US 6844180.			781	CTAGTGGAAACATAGTTGTCCTCAACATAGCGATCGATCCAAATTTATGCCCGCTGATCCA 840
ACCESSION	AR631288			781	CTAGTGGAAACATAGTTGTCCTCAACATAGCGATCGATCCAAATTTATGCCCGCTGATCCA 840
VERSION	AR631288.1	GI:59770936		841	GACTGGATTGCTTTTATACATAGCAACGATATTTGGATATCTAACAATCTTATGAAGAA 900
KEYWORDS	Unknown.			841	GACTGGATTGCTTTTATACATAGCAACGATATTTGGATATCTAACAATCTTATGAAGAA 900
SOURCE	Unknown.			901	GAAAGGAGACTCCTATATGTGCAACAATGAGCTAGCCAAACATGGAAGAAAGATGCCAGATCA 960
ORGANISM	Unclassified.			901	GAAAGGAGACTCCTATATGTGCAACAATGAGCTAGCCAAACATGGAAGAAAGATGCCAGATCA 960
REFERENCE	1 (bases 1 to 4685)			961	GCTGGAGTCGCTACCTTTGTTTCTCCAAAGAAATTTGATAGATATCTGGCTATTGGTGG 1020
AUTHORS	Oi, S., Akinsanya, K.O., Riviere, P.J.M. and Junien, J.-L.			961	GCTGGAGTCGCTACCTTTGTTTCTCCAAAGAAATTTGATAGATATCTGGCTATTGGTGG 1020
TITLE	Serine protease genes related to DPPIV			1021	TGTCANAAAGCTGAAACAACTCCAGTGGTGTAAATTTCTTAGAAATCTTATATGAAGAA 1080
JOURNAL	Patent: US 6844180-A 22 18-JAN-2005;			1021	TGTCANAAAGCTGAAACAACTCCAGTGGTGTAAATTTCTTAGAAATCTTATATGAAGAA 1080
	Ferring BV,;			1081	AATGATGATCTGAGCTGGAAATTTATCATGTTACATCCCTATGTTGGAAACAAGGAGG 1140
FEATURES	Location/Qualifiers			1081	AATGATGATCTGAGCTGGAAATTTATCATGTTACATCCCTATGTTGGAAACAAGGAGG 1140
source	1. 4685			1141	GCAGATTCATCCGTTTATCTTAAACAGGATCAGCAAAATCTTAAAGTCACTTTTAGATG 1200
ORIGIN	/organism="unknown"			1141	GCAGATTCATCCGTTTATCTTAAACAGGATCAGCAAAATCTTAAAGTCACTTTTAGATG 1200
	/mol_type="genomic DNA"			1201	TCAGAAATTAATGATGCTGAAAGGAAGATCATAGATGTCATAGATATAGGAACATTAAT 1260
Query Match	90.5%; Score 2824.4; DB 6; Length 4685;			1201	TCAGAAATTAATGATGCTGAAAGGAAGATCATAGATGTCATAGATATAGGAACATTAAT 1260
Best Local Similarity	95.4%; Pred. No. 0;			1261	CAACCTTTTGAGATCTTATTTGAAAGGAGTTGAATATATTCGAGAGCTGATGGACTCCT 1320
Matches 2977; Conservative 0; Mismatches 1; Indels 142; Gaps 1;				1261	CAACCTTTTGAGATCTTATTTGAAAGGAGTTGAATATATTCGAGAGCTGATGGACTCCT 1320
Qy	1	AAGTGCTTAAAGCTCCGAGGCCAAGGCGCTGTCTACTGCGCGCGCTTCTTAGTGGCG 60		1321	GAGGGAATAATATGCTTGGTCCATCCTACTAGATCGCTCCAGACTCGCTACAGATAGTG 1380
Db	1	AAGTGCTTAAAGCTCCGAGGCCAAGGCGCTGTCTACTGCGCGCGCTTCTTAGTGGCG 60		1321	GAGGGAATAATATGCTTGGTCCATCCTACTAGATCGCTCCAGACTCGCTACAGATAGTG 1380
Qy	61	CGTTCCGCGCTGGGTTGTCAACCGCGCCCGCGGAGGAAAGCACTGCAACAGGACCG 120		1381	TTGATCTCACCTGAATTTATTTATCCAGTGAAGATGATGTTATGGAAAGGCGAGAGACTC 1440
Db	61	CGTTCCGCGCTGGGTTGTCAACCGCGCCCGCGGAGGAAAGCACTGCAACAGGACCG 120		1381	TTGATCTCACCTGAATTTATTTATCCAGTGAAGATGATGTTATGGAAAGGCGAGAGACTC 1440
Qy	121	GAGTGGAGCGCGCAGCATGAAGCGCGCAGGCGCGCTCCATAGCGCAGCTCGGAGCGG 180		1441	ATTGAGTCAGTCGCTGATCTGTGACGCCACTTAATTTCTATGAGAAACAACAGACATC 1500
Db	121	GAGTGGAGCGCGCAGCATGAAGCGCGCAGGCGCGCTCCATAGCGCAGCTCGGAGCGG 180		1441	ATTGAGTCAGTCGCTGATCTGTGACGCCACTTAATTTCTATGAGAAACAACAGACATC 1500
Qy	181	TCGGGCGGGGCGGGGGAAGGAAATGCAACATGGCAGCAGCAATGGAACAGAAACAG 240		1501	TGATAAATATCCATGACATCTTTCATGTTTCCCAAGTCAAGAGGAAATTTGAG 1560
Db	181	TCGGGCGGGGCGGGGGAAGGAAATGCAACATGGCAGCAGCAATGGAACAGAAACAG 240		1501	TGATAAATATCCATGACATCTTTCATGTTTCCCAAGTCAAGAGGAAATTTGAG 1560
Qy	241	CTGGGTGTGAGATTTGAACTGCGGACTGTGAGGAGATATTCGAATCAGAGATCGG 300		1561	TTTATTTTTCCTCTGAAATGCAAAACAGGTTTCCGTCATTTATACAAATTTACATCTATT 1620
Db	241	CTGGGTGTGAGATTTGAACTGCGGACTGTGAGGAGATATTCGAATCAGAGATCGG 300		1561	TTTATTTTTCCTCTGAAATGCAAAACAGGTTTCCGTCATTTATACAAATTTACATCTATT 1620
Qy	301	CCTAAATTGGAGCCCTTTTATTTGAGGGGTATTCCTGGAGTCAGCTTAAAGCTGCTT 360		1621	TTAAAGGAAGCAAAATATAACGATCCAGTGGTGGCTGCTGCTCCAAAGTATTCAAG 1680
Db	301	CCTAAATTGGAGCCCTTTTATTTGAGGGGTATTCCTGGAGTCAGCTTAAAGCTGCTT 360		1621	TTAAAGGAAGCAAAATATAACGATCCAGTGGTGGCTGCTGCTCCAAAGTATTCAAG 1680
Qy	361	GCCGATACAGAAAAATATCATGGCTACATGATGGCTAAGGCAACAATGATTTTCAATGTTT 420		1681	TGTCTCTATCAAAAGAGGAGATAGCAATTTACCAAGTGGTGAATGGGAAGTCTTTGGCCGCAT 1740
Db	361	GCCGATACAGAAAAATATCATGGCTACATGATGGCTAAGGCAACAATGATTTTCAATGTTT 420		1681	TGTCTCTATCAAAAGAGGAGATAGCAATTTACCAAGTGGTGAATGGGAAGTCTTTGGCCGCAT 1740
Qy	421	GTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGAAATCTATTACCTTGCATGCT 480			
Db	421	GTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGAAATCTATTACCTTGCATGCT 480			
Qy	481	GGTGAGAACAGAGAAAAATACACTGTTTATTTCTGAAATTCCTCAAACTATCAATAGACA 540			
Db	481	GGTGAGAACAGAGAAAAATACACTGTTTATTTCTGAAATTCCTCAAACTATCAATAGACA 540			
Qy	541	GCAGTCTTAATGCTCTCTTGAAGCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT 600			
Db	541	GCAGTCTTAATGCTCTCTTGAAGCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT 600			
Qy	601	CGAATGTATCTCGAGAGAGAACTATTAGAGAGAAAGCAATGGAAACAGTCGGA 660			


```
QY 1741 GGATCTAATATCCAAAGTTGATGAAGTCAGAGGCTGGTATATATTTTGAAGGCCAACAGAC 1800
Db 1741 GGATCTAATATCCAAAGTTGATGAAGTCAGAGGCTGGTATATATTTTGAAGGCCAACAGAC 1800
QY 1801 TCCCTTTAGAGCATACCTGTAGTACGTTAGTACGTTAGTAAATCTGGAGAGGTCAAGG 1860
Db 1801 TCCCTTTAGAGCATACCTGTAGTACGTTAGTACGTTAGTAAATCTGGAGAGGTCAAGG 1860
QY 1861 CTGACTGACCGTGGCTACTCACATCTCTGCTGATCAGTCAGCACTGTGACTCTTTTATA 1920
Db 1861 CTGACTGACCGTGGCTACTCACATCTCTGCTGATCAGTCAGCACTGTGACTCTTTTATA 1920
QY 1921 AGTAAGTATAGTAACCAAGAGAAATCCACTGTGTGTCCTTTTACAAGCTATCAAGTCCT 1980
Db 1921 AGTAAGTATAGTAACCAAGAGAAATCCACTGTGTGTCCTTTTACAAGCTATCAAGTCCT 1980
QY 1981 GAAGATGACCACTTGCACAAACAAAGGAATTTGGGCCACCATTTGGATTCAGAGGT 2040
Db 1981 GAAGATGACCACTTGCACAAACAAAGGAATTTGGGCCACCATTTGGATTCAGAGGT 2040
QY 2041 CCTCTTCTGACTATCTCCTCCAGAAATTTTCTCTTTTGAAGTACTACTGATTTTACA 2100
Db 2035 ----- 2034
QY 2101 TTGTATGGGATGCTCTACAAGCCTCATGATCTACAGCCTTGGAAGAAATATCTTACTGTG 2160
Db 2035 ----- 2034
QY 2161 CTGTTTCATATATGTTGTCCTCAGGTGTCAGTTGGTGATATATCGGTTTAAAGGATCAAG 2220
Db 2035 -----GTCCTCAGGTGTCAGTTGGTGATATATCGGTTTAAAGGATCAAG 2078
QY 2221 TATTTCCGTTGATATACCTAGCTCTCTAGTTATGTTGTTAGTATAGACACAAAG 2280
Db 2079 TATTTCCGTTGATATACCTAGCTCTCTAGTTATGTTGTTAGTATAGACACAAAG 2138
QY 2281 GGATCCTGTCTACCGAGGCTTAAATTTGAAGGCGCTTTAAATATATAAATGGTCAATA 2340
Db 2139 GGATCCTGTCTACCGAGGCTTAAATTTGAAGGCGCTTTAAATATATAAATGGTCAATA 2198
QY 2341 GAAATTCAGCATCAGGTGGAAGGACTCCAATATCTAGTTCTTCGATATGATTTCAATGAC 2400
Db 2199 GAAATTCAGCATCAGGTGGAAGGACTCCAATATCTAGTTCTTCGATATGATTTCAATGAC 2258
QY 2401 TTAGATCTGTGGGCATCCAGCGTGTCTATGAGGATACCTCTCCCTGATGGCATT 2460
Db 2259 TTAGATCTGTGGGCATCCAGCGTGTCTATGAGGATACCTCTCCCTGATGGCATT 2318
QY 2461 ATGCAGAGGTCAGATATCTTCAGGTTGCTTATGTTGGGCGCCCAAGTCACTCTGTGGATC 2520
Db 2319 ATGCAGAGGTCAGATATCTTCAGGTTGCTTATGTTGGGCGCCCAAGTCACTCTGTGGATC 2378
QY 2521 TTCTATGATACAGGATACAGGAAAGTATATGGGTCACTTCGATGATGATGATGATGATG 2580
Db 2379 TTCTATGATACAGGATACAGGAAAGTATATGGGTCACTTCGATGATGATGATGATGATG 2438
QY 2581 TATTTACTAGGATCTGTGGCCATGCAAGCAGAAAGTTTCCCTCTGCAACCAATCGTTTA 2640
Db 2439 TATTTACTAGGATCTGTGGCCATGCAAGCAGAAAGTTTCCCTCTGCAACCAATCGTTTA 2498
QY 2641 CTGCTCTTACATGTTTCTGGATGAGAAATGTCATTTTGCATACATACCATATATCTG 2700
Db 2499 CTGCTCTTACATGTTTCTGGATGAGAAATGTCATTTTGCATACATACCATATATCTG 2558
QY 2701 AGTTTTTTTGTAGTGGGCTGGAAAGCCTATGATTTACAGATCTATCTCTCAGGAGACAC 2760
Db 2559 AGTTTTTTTGTAGTGGGCTGGAAAGCCTATGATTTACAGATCTATCTCTCAGGAGACAC 2618
QY 2761 AGCATAGAGTTCTGAAATCGGGAGAACATTATGAATCTGATCTTTTGTGACATACCTTCA 2820
Db 2619 AGCATAGAGTTCTGAAATCGGGAGAACATTATGAATCTGATCTTTTGTGACATACCTTCA 2678
```

```
QY 2821 GAAAACTTGGATCAGTATTCGTCTTAAAGTGATATATATTTTGACCTGTGTAGAAC 2880
Db 2679 GAAAACTTGGATCAGTATTCGTCTTAAAGTGATATATATTTTGACCTGTGTAGAAC 2738
QY 2881 TCTCTGTATACACTGCTGCTATTTAAACCAATGAGGAGGTTTAAATCAACAGAAAAACAGA 2940
Db 2739 TCTCTGTATACACTGCTGCTATTTAAACCAATGAGGAGGTTTAAATCAACAGAAAAACAGA 2798
QY 2941 ATTGATCATCATATTTTGATACCTGCGATGTAACATCTACTCTCGAAAAATTAATTTGGTG 3000
Db 2799 ATTGATCATCATATTTTGATACCTGCGATGTAACATCTACTCTCGAAAAATTAATTTGGTG 2858
QY 3001 CCATGAGGCGTCTACGGTTTGTGTAGTAAATCTTAACTTAACTTAACTTAACTTAACTTAACT 3060
Db 2859 CCATGAGGCGTCTACGGTTTGTGTAGTAAATCTTAACTTAACTTAACTTAACTTAACTTAACT 2918
QY 3061 TCAATGATACATATTTCTGAGAGACCCAGCAATACCAATAAGAAATTAATAAAAAA 3120
Db 2919 TCAATGATACATATTTCTGAGAGACCCAGCAATACCAATAAGAAATTAATAAAAAA 2978
RESULT 9
AX608745 4685 bp DNA linear PAT 17-FEB-2003
LOCUS AX608745 Sequence 22 from Patent WO0231134.
DEFINITION AX608745
ACCESSION AX608745.1 GI:28404308
VERSION
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE
1 Qi, S., Akinsanya, K.O., Riviere, P.J. and Junien, J.L.
AUTHORS Novel serine protease genes related to dplv
TITLE Patent: WO 0231134-A 22 18-APR-2002;
JOURNAL Ferring BV (NL)
FEATURES
Location/Qualifiers
source 1..4685
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 90.5%; Score 2824.4; DB 6; Length 4685;
Best Local Similarity 95.4%; Pred. No. 0;
Matches 2977; Conservative 0; Mismatches 1; Indels 142; Gaps 1;
QY 1 AAGTGCTAAAGCCTCCGAGGCCAAGCCGCTGCTACTGCGCCGCTGCTTCTTAGTCCG 60
Db 1 AAGTGCTAAAGCCTCCGAGGCCAAGCCGCTGCTACTGCGCCGCTGCTTCTTAGTCCG 60
QY 61 CGTTGCGCGCTGTTGTTGTCACCGCGCGCGCGCGAGGAGCCACTCCACCCAGGACCG 120
Db 61 CGTTGCGCGCTGTTGTTGTCACCGCGCGCGCGCGAGGAGCCACTCCACCCAGGACCG 120
QY 121 GAGTGGAGGCGCGCGAGCATGAAGCGCGCGCGCGCTCCATAGCGACGCTCGGGA 180
Db 121 GAGTGGAGGCGCGCGAGCATGAAGCGCGCGCGCGCTCCATAGCGACGCTCGGGA 180
QY 181 TCCGCGCGCGCGCGGAGGAAATGCAACATGCGAGCAGCAATGGAACAGAACAG 240
Db 181 TCCGCGCGCGCGCGGAGGAAATGCAACATGCGAGCAGCAATGGAACAGAACAG 240
QY 241 CTGGGTGTTGATATTTTGAACCTGCGGACTGTGAGGAGCAATATTGAATCAGCATCGG 300
Db 241 CTGGGTGTTGATATTTTGAACCTGCGGACTGTGAGGAGCAATATTGAATCAGCATCGG 300
QY 301 CCTAAATTTGGAGCCTTTTATGTTGAGCGGTATTTCTGAGTCACTTAAAGCTGCTT 360
Db 301 CCTAAATTTGGAGCCTTTTATGTTGAGCGGTATTTCTGAGTCACTTAAAGCTGCTT 360
```


Qy	361	GCCGATACAGAAAATATCATGCTCATGCTAAGGACACACATGATTTTCATGTTT	420	Db	1441	ATTGAGTCAGTGGCTGATTTCTGTGACGCCATTAATTTCTATGAAAGAACACAGACATC	1500
Db	361	GCCGATACAGAAAATATCATGCTCATGCTAAGGACACACATGATTTTCATGTTT	420	Qy	1501	TGGATAAATATCCATGACATCTTTTCATGTTTTTCCCAAGTCACGAGAGGAAATTCAG	1560
Qy	421	GTGAGAGGAATGATCCAGATGGACCTCATTCACAGACAGAAATCTATTACCTTGGCCATGCT	480	Db	1501	TGGATAAATATCCATGACATCTTTTCATGTTTTTCCCAAGTCACGAGAGGAAATTCAG	1560
Db	421	GTGAGAGGAATGATCCAGATGGACCTCATTCACAGACAGAAATCTATTACCTTGGCCATGCT	480	Qy	1561	TTTATTTTTGCTCTGAAATGCAAAAACAGGTTTCCGTCATTTTATACAAAATATACATCTATT	1620
Qy	481	GCTGAGAACAGAGAAAATACACTGTTTTTATCTTGAATTTCCCAAACTATCAATAGAGCA	540	Db	1561	TTTATTTTTGCTCTGAAATGCAAAAACAGGTTTCCGTCATTTTATACAAAATATACATCTATT	1620
Db	481	GCTGAGAACAGAGAAAATACACTGTTTTTATCTTGAATTTCCCAAACTATCAATAGAGCA	540	Qy	1621	TTAAAGGAAACAAATATAAACGATCCAGTGGTGGCTGCTGCTCCAAAGTATTTCAAG	1680
Qy	541	GCAGCTTTAAATGCTCTCTTGAAGCCTCTTTTGGATCTTTTTCAGGCAACACTGACATAT	600	Db	1621	TTAAAGGAAACAAATATAAACGATCCAGTGGTGGCTGCTGCTCCAAAGTATTTCAAG	1680
Db	541	GCAGCTTTAAATGCTCTCTTGAAGCCTCTTTTGGATCTTTTTCAGGCAACACTGACATAT	600	Qy	1681	TGTCCTATCAAAAGAGGAGATAGCAATTAACAGTGGTGAATGGGAAGTTCTTGGCCGGCAT	1740
Qy	601	GGAATGTATTCGAGAGAAAGAAACTATTAAGAGAAAGAAACGCATTTGGAAACAGTCGGA	660	Db	1681	TGTCCTATCAAAAGAGGAGATAGCAATTAACAGTGGTGAATGGGAAGTTCTTGGCCGGCAT	1740
Db	601	GGAATGTATTCGAGAGAAAGAAACTATTAAGAGAAAGAAACGCATTTGGAAACAGTCGGA	660	Qy	1741	GGATCTAATATCCAAGTTGATGAAGTCAGAAAGCTGGTATATTTTGAAGGCCACCAAGAC	1800
Qy	661	ATTGCTTTCTTACGATTAATCACCAAGGAAGTGGAACTTTCTGTTTCAAGCCGGTAGTGA	720	Db	1741	GGATCTAATATCCAAGTTGATGAAGTCAGAAAGCTGGTATATTTTGAAGGCCACCAAGAC	1800
Db	661	ATTGCTTTCTTACGATTAATCACCAAGGAAGTGGAACTTTCTGTTTCAAGCCGGTAGTGA	720	Qy	1801	TCGCCCTTTAGAGCATCACCTGTAGTCAGTTACGTAAATCCTCGAGAGGTGACAAGG	1860
Qy	721	ATTATATCACTTAAAGATGAGAGGCCCAAGGATTTTACGCAACAACTTTAAAGGCCAAT	780	Db	1801	TCGCCCTTTAGAGCATCACCTGTAGTCAGTTACGTAAATCCTCGAGAGGTGACAAGG	1860
Db	721	ATTATATCACTTAAAGATGAGAGGCCCAAGGATTTTACGCAACAACTTTAAAGGCCAAT	780	Qy	1861	CTGACTGACCGTGGCTACTCACATTTCTGCTGCATCAGTCAGCACGTGACTTCTTTATA	1920
Qy	781	CTAGTGGAAACTAGTTGTCCCAACATACGGATGGATCCAAAATTTATGCCCCGCTGATCCA	840	Db	1861	CTGACTGACCGTGGCTACTCACATTTCTGCTGCATCAGTCAGCACGTGACTTCTTTATA	1920
Db	781	CTAGTGGAAACTAGTTGTCCCAACATACGGATGGATCCAAAATTTATGCCCCGCTGATCCA	840	Qy	1921	AGTAAAGTATAGTAAACAGAGAAATCCACACTGTGTGTCCTTTTACAAAGCTATCAAGTCTCT	1980
Qy	841	GACTGGATTTGCTTTTATACATAGCAACGATATTTGGATATCTTAACATCGTAAACAGAGAA	900	Db	1921	AGTAAAGTATAGTAAACAGAGAAATCCACACTGTGTGTCCTTTTACAAAGCTATCAAGTCTCT	1980
Db	841	GACTGGATTTGCTTTTATACATAGCAACGATATTTGGATATCTTAACATCGTAAACAGAGAA	900	Qy	1981	GAAGATGACCCAACTTGGCAAAACAAAGAAATTTTGGCCGACCATTTTGGATTTCA	2040
Qy	901	GAAGAGACTCAGTTATGTCACAAATGATGATGATGATGATGATGATGATGATGATGATGAT	960	Db	1981	GAAGATGACCCAACTTGGCAAAACAAAGAAATTTTGGCCGACCATTTTGGATTTCA	2040
Db	901	GAAGAGACTCAGTTATGTCACAAATGATGATGATGATGATGATGATGATGATGATGATGAT	960	Qy	2041	CCTCTTCTGACTATATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGATTTACA	2100
Qy	961	GCTGGAGTGGCTACCTTTGCTCCAGAGAAATTTGATAGATATTTCTGGCTATTTGGTGG	1020	Db	2035	-----	2034
Db	961	GCTGGAGTGGCTACCTTTGCTCCAGAGAAATTTGATAGATATTTCTGGCTATTTGGTGG	1020	Qy	2101	TTGTATGGGATGCTCTACAAAGCTCATGATCTACAGCCTGGAAAGAAATATCCTACTGTG	2160
Qy	1021	TGTCCAAAAGCTGAAACAACTCCAGTGGTGGTAAATTTCTTAGAATTTCTATATGAAGAA	1080	Db	2035	-----	2034
Db	1021	TGTCCAAAAGCTGAAACAACTCCAGTGGTGGTAAATTTCTTAGAATTTCTATATGAAGAA	1080	Qy	2161	CTGTTTCATATATGTTGGTCTCCTCAGGTGCAAGTTGGTGAATTAATCGGTTTAAAGGAGTCAAG	2220
Qy	1081	AATGATGAATCTGAGGTGAAATTTATCATGTTACATCCCTATGTTGGAAAACAGGAGG	1140	Db	2035	-----GTCCTCAGGTGCAAGTTGGTGAATTAATCGGTTTAAAGGAGTCAAG	2078
Db	1081	AATGATGAATCTGAGGTGAAATTTATCATGTTACATCCCTATGTTGGAAAACAGGAGG	1140	Qy	2221	TATTTCCGCTTGAATACCTCTAGCTCTCTAGCTTATGTTGGTGTAGTAGTAGACAACAGG	2280
Qy	1141	GCAGATTCATTCGTTTATCCTTAAACAGGTACAGCAAAATCCTAAAGTCACTTTTAAAGTG	1200	Db	2079	TATTTCCGCTTGAATACCTCTAGCTTATGTTGGTGTAGTAGTAGACAACAGG	2138
Db	1141	GCAGATTCATTCGTTTATCCTTAAACAGGTACAGCAAAATCCTAAAGTCACTTTTAAAGTG	1200	Qy	2281	GGATCCTGTACCCGAGGGCTTAAATTTGAAGCGGCTTTTAAATATAAAATGGGTCAAAATA	2340
Qy	1201	TCAGAAATATGATGATGCTGAAGGAAGGATCATAGATGTCTAGAGATAAGGAACATAAT	1260	Db	2139	GGATCCTGTACCCGAGGGCTTAAATTTGAAGCGGCTTTTAAATATAAAATGGGTCAAAATA	2198
Db	1201	TCAGAAATATGATGATGCTGAAGGAAGGATCATAGATGTCTAGAGATAAGGAACATAAT	1260	Qy	2341	GAATTTGACGATCAGGTGGAAGGATCCCAATATCTAGCTTCTCGATATGATTTTCATTGAC	2400
Qy	1261	CAACCTTTTACAGATTTCTTTTGAAGGATTTGAATATATTTGCCAGAGCTGGATGGACTCCT	1320	Db	2199	GAATTTGACGATCAGGTGGAAGGATCCCAATATCTAGCTTCTCGATATGATTTTCATTGAC	2258
Db	1261	CAACCTTTTACAGATTTCTTTTGAAGGATTTGAATATATTTGCCAGAGCTGGATGGACTCCT	1320	Qy	2401	TTAGATCGTGGGATCCACGGCTGCTTATGCTGAGGATACCTCTCCCTGATGATTA	2460
Qy	1321	GAGGAAATATGCTTGGTCCATCTCTACTAGATCGTCCAGACTCGCTACAGATAGTG	1380	Db	2259	TTAGATCGTGGGATCCACGGCTGCTTATGCTGAGGATACCTCTCCCTGATGATTA	2318
Db	1321	GAGGAAATATGCTTGGTCCATCTCTACTAGATCGTCCAGACTCGCTACAGATAGTG	1380	Qy	2461	ATGAGAGGTGAGATATCTTTCAAGGTGCTTATGCTGGGGCCGACGTCACCTCTGGATC	2520
Qy	1381	TTGATCTCACTGAAATTTTATCCAGTAGAGATGATGTTATGGAAGGACAGAGACTC	1440	Db	2319	ATGAGAGGTGAGATATCTTTCAAGGTGCTTATGCTGGGGCCGACGTCACCTCTGGATC	2378
Db	1381	TTGATCTCACTGAAATTTTATCCAGTAGAGATGATGTTATGGAAGGACAGAGACTC	1440	Qy	2521	TTCTATGATACAGGATACAGGAACTTATATGGGTACCTCTGACCAAGTAAGACAGGGC	2580
Qy	1441	ATTGAGTCAGTGGCTGATTTCTGTGACGCCACTAAATTTATCTATGAAAGAACACAGACATC	1500				

Db 2379 TTCTATGATACAGGATACAGGAACGTTATATGGGTCAACCTGACCAAGATGAAACAGGCG 2438

Qy 2581 TATTACTTAGGATCTGTGGCCATGCAAGCAGAGAAAGTTCCCTCTGAAACAAATCGTTTA 2640

Db 2439 TATTACTTAGGATCTGTGGCCATGCAAGCAGAGAAAGTTCCCTCTGAAACAAATCGTTTA 2498

Qy 2641 CAGCTCTTACATGGTTTCTCTGGATGAGAAAGTGTCCATTTTGCACATCCAGTATATTACTG 2700

Db 2499 CTGCTCTTACATGGTTTCTCTGGATGAGAAAGTGTCCATTTTGCACATCCAGTATATTACTG 2558

Qy 2701 AGTTTTTTAGTGAGGCTGGAAGCCATATGATTTACAGATCTATCTCTCAGGAGAGACAC 2760

Db 2559 AGTTTTTTAGTGAGGCTGGAAGCCATATGATTTACAGATCTATCTCTCAGGAGAGACAC 2618

Qy 2761 AGCATAAGAGTTCTCTGAATCGGGAGAACATTTATGAATGCAATCTTTTGCACATCTTCAA 2820

Db 2619 AGCATAGAGTTCTCTGAATCGGGAGAACATTTATGAATGCAATCTTTTGCACATCTTCAA 2678

Qy 2821 GAAACCTTTGGATCAGTATTTGCTGCTCTAAAGTGATATAATTTTGAATGCTGTGTAGAAC 2880

Db 2679 GAAACCTTTGGATCAGTATTTGCTGCTCTAAAGTGATATAATTTTGAATGCTGTGTAGAAC 2738

Qy 2881 TCTCTGGTATACATCGGCTATTTAAACCAATGAGAGGTTTAAATCAACAGAAACACAGA 2940

Db 2739 TCTCTGGTATACATCGGCTATTTAAACCAATGAGAGGTTTAAATCAACAGAAACACAGA 2798

Qy 2941 ATTGATCATCAATTTTGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3000

Db 2799 ATTGATCATCAATTTTGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2858

Qy 3001 CCATGCGGGGCTACGGTTTGGTGTAGTAACTTAATTAACCTTAACCCCATGCTCAAAA 3060

Db 2859 CCATGCGGGGCTACGGTTTGGTGTAGTAACTTAATTAACCTTAACCCCATGCTCAAAA 2918

Qy 3061 TCAATGATACATATTTCTGAGAGCCAGCAATACCAATGATTAATTAATTAATTAATTAATTA 3120

Db 2919 TCAATGATACATATTTCTGAGAGCCAGCAATACCAATGATTAATTAATTAATTAATTAATTA 2978

RESULT 10

AR631287

LOCUS AR631287 4676 bp DNA linear PAT 14-FEB-2005

DEFINITION Sequence 20 from patent US 6844180.

ACCESSION AR631287

VERSION AR631287.1 GI:59770934

KEYWORDS

SOURCE Unknown.

ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 4676)

AUTHORS Oi,S., Akinsanya,K.O., Riviere,P.J.M. and Junien,J.-L.

TITLE Serine protease genes related to DPPIV

JOURNAL Patent: US 6844180-A 20 18-JAN-2005;

Ferring BV,;

NLX;

FEATURES

source

1..4676

/organism="unknown"

/mol_type="genomic DNA"

ORIGIN

Query Match 89.9%; Score 2806.4; DB 6; Length 4676;

Best Local Similarity 95.1%; Pred. No. 0;

Matches 2968; Conservative 0; Mismatches 1; Indels 151; Gaps 1;

Qy 1 AAGTGCTAAAGCTCCGAGGCCAAGCGCGCTGCTACTGCGCGCGCTGCTTCTTAGTGCCG 60

Db 1 AAGTGCTAAAGCTCCGAGGCCAAGCGCGCTGCTACTGCGCGCGCTGCTTCTTAGTGCCG 60

Qy 61 CGTTCCGCGCTGGGTGTGTCACCG 120

Db 61 CGTTCCGCGCTGGGTGTGTCACCG 120

Qy 121 GAGTCGAGCGCGCAGCATGAAGCGCGCAGGCGCGCTCCATAGCGCACGTCGGGACGG 180

Db 121 GAGTCGAGCGCGCAGCATGAAGCGCGCAGGCGCGCTCCATAGCGCACGTCGGGACGG 180

Qy 181 TCCGGGCGGGCGCGGGGAGGAAATGCAACATGCGCAGCAGCAATGGAACAGAACAG 240

Db 181 TCCGGGCGGGCGCGGGGAGGAAATGCAACATGCGCAGCAGCAATGGAACAGAACAG 240

Qy 241 CTGGGTGTTGAGATATTGAAACTCGGACCTGTGAGGAGATATTGATCAGAGATCGG 300

Db 241 CTGGGTGTTGAGATATTGAAACTCGGACCTGTGAGGAGATATTGATCAGAGATCGG 300

Qy 301 CCTAAATTCGAGCTTTTATGTCGAGCGGTATTCCTCGAGTCAGCTTTAAAGAGCTGCT 360

Db 301 CCTAAATTCGAGCTTTTATGTCGAGCGGTATTCCTCGAGTCAGCTTTAAAGAGCTGCT 360

Qy 361 GCCGATACAGAGAAATATCATGCTACATGATGGCTAAGGCCACACATGATTTTATGTTT 420

Db 361 GCCGATACAGAGAAATATCATGCTACATGATGGCTAAGGCCACACATGATTTTATGTTT 420

Qy 421 GTGAAGAGGAATGATCCAGATGGAACCTCATTCAGACAGAACTTATACCTTGGCATGCT 480

Db 421 GTGAAGAGGAATGATCCAGATGGAACCTCATTCAGACAGAACTTATACCTTGGCATGCT 480

Qy 481 GGTGAGACAGAGAAATACACTGTTTATTCGAAATTCCTGAAATTCCTGAAATTCCTGAAATTC 540

Db 481 GGTGAGACAGAGAAATACACTGTTTATTCGAAATTCCTGAAATTCCTGAAATTCCTGAAATTC 540

Qy 541 GCAGTCTTAAATGCTCTCTTGGAGGCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT 600

Db 541 GCAGTCTTAAATGCTCTCTTGGAGGCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT 600

Qy 601 GGAATGTATTCTCGAAGAGAACTATTTAAGAGAAAGAAACGCAATTTGGAACAGTCGGA 660

Db 601 GGAATGTATTCTCGAAGAGAACTATTTAAGAGAAAGAAACGCAATTTGGAACAGTCGGA 660

Qy 661 ATTGCTTCTTAAGATATCCAGAGAGTGAACATTTCTGTTTCAAGCGGTAGTGA 720

Db 661 ATTGCTTCTTAAGATATCCAGAGAGTGAACATTTCTGTTTCAAGCGGTAGTGA 720

Qy 721 ATTTATCAGTAAAGATGAGGCGGCACAGGATTTTACGCAACACTTTAAGGCCCAAT 780

Db 721 ATTTATCAGTAAAGATGAGGCGGCACAGGATTTTACGCAACACTTTAAGGCCCAAT 780

Qy 781 CTAGTGGAACCTAGTTTGTCCCAACATACGATGGAATCCAAATTTATGCCCCGCTGATCCA 840

Db 781 CTAGTGGAACCTAGTTTGTCCCAACATACGATGGAATCCAAATTTATGCCCCGCTGATCCA 840

Qy 841 GACTGATGCTTTTATACATAGCAACGATTTTGGATATCTAACATCTGAACACAGAA 900

Db 841 GACTGATGCTTTTATACATAGCAACGATTTTGGATATCTAACATCTGAACACAGAA 900

Qy 901 GAAAGGAGCTCACTTATGTCACATGAGCTAGCCACATGGAAGAGATGCCAGATCA 960

Db 901 GAAAGGAGCTCACTTATGTCACATGAGCTAGCCACATGGAAGAGATGCCAGATCA 960

Qy 961 GCTGAGTCTGCTACCTTTGTTCTCCAAAGAAATTTGATAGATTTCTGGCTATTGGTGG 1020

Db 961 GCTGAGTCTGCTACCTTTGTTCTCCAAAGAAATTTGATAGATTTCTGGCTATTGGTGG 1020

Qy 1021 TGTCCAAAGCTGAAACAACTCCAGTGGTGAATTTCTTGAATTTCTATATGAAGAA 1080

Db 1021 TGTCCAAAGCTGAAACAACTCCAGTGGTGAATTTCTTGAATTTCTATATGAAGAA 1080

Qy 1081 AATGATGAATCTGAGTGGAAATTTATTCATGTTATACCTCCCTATGTTTGGAAACAGGAGG 1140

Db 1081 AATGATGAATCTGAGTGGAAATTTATTCATGTTATACCTCCCTATGTTTGGAAACAGGAGG 1140

Qy 1141 GCAGATTCATTCTGTTTATCTTAAACAGGTACAGCAAAATCTTAAAGTCACTTTTAAAGATG 1200

Db 1141 GCAGATTCATTCTGTTTATCTTAAACAGGTACAGCAAAATCTTAAAGTCACTTTTAAAGATG 1200

Qy 1201 TCAGAAATATGATTTGATGCTGAAGAGGATCATAGATGTCATAGATAAGGAACATAAT 1260

1201	Db	TCAGAAATAATGATTTGATGCTGAAGGAAGGATCATAGATGTCATAGATAGGAACATAAT	1260
1261	Qy	CAACCTTTTGAGATTCTATTTGGAAGGAGTTGAATATATTTGCCAGAGCTGGATGGACTCCT	1320
1261	Db	CAACCTTTTGAGATTCTATTTGGAAGGAGTTGAATATATTTGCCAGAGCTGGATGGACTCCT	1320
1321	Qy	GAGGGAATAATGCTTTGGTCCATCTCTACTAGATCGCTCCAGACTCGCTACAGATAGTG	1380
1321	Db	GAGGGAATAATGCTTTGGTCCATCTCTACTAGATCGCTCCAGACTCGCTACAGATAGTG	1380
1381	Qy	TTGATCTCACCTGGAATTATTTATCCCAAGTAGAAGATGATGTTATGSGAAAGGCAGAGACTC	1440
1381	Db	TTGATCTCACCTGGAATTATTTATCCCAAGTAGAAGATGATGTTATGSGAAAGGCAGAGACTC	1440
1441	Qy	ATTGAGTCAGTGCCTGATCTTGTCAGGCCATAATTTATCTATGAAGAAACAACAGACATC	1500
1441	Db	ATTGAGTCAGTGCCTGATCTTGTCAGGCCATAATTTATCTATGAAGAAACAACAGACATC	1500
1501	Qy	TGGATAAATATCCATGACATCTTTCATGTTTTCCCAAGTCACGAGAGGGAATTTGAG	1560
1501	Db	TGGATAAATATCCATGACATCTTTCATGTTTTCCCAAGTCACGAGAGGGAATTTGAG	1560
1561	Qy	TTTATTTTGGCTCTGAAATGCAAAACAGGTTTCCGTCAATTTATACAAATTTACATCTATT	1620
1561	Db	TTTATTTTGGCTCTGAAATGCAAAACAGGTTTCCGTCAATTTATACAAATTTACATCTATT	1620
1621	Qy	TTAAAGGAAAGCAATATAAACGATCCAGTGTGGGCTGCCGTCTCAAGTGATTTCAAG	1680
1621	Db	TTAAAGGAAAGCAATATAAACGATCCAGTGTGGGCTGCCGTCTCAAGTGATTTCAAG	1680
1681	Qy	TGTCCTATCAAGAGGAGATAGCAATTTACCAAGTGTGGAATGGGAAGTTCTTTGGCGGCGAT	1740
1681	Db	TGTCCTATCAAGAGGAGATAGCAATTTACCAAGTGTGGAATGGGAAGTTCTTTGGCGGCGAT	1740
1741	Qy	GGATCTAATATCCAAGTTGATGAAGTCAGAAGGCTGGTATATTTTTGAAGGCACCAAGAC	1800
1741	Db	GGATCTAATATCCAAGTTGATGAAGTCAGAAGGCTGGTATATTTTTGAAGGCACCAAGAC	1800
1801	Qy	TCCTCTTAGAGCATCACCTGTACGTAGTCAGTTACGTAAATCTGGAGAGGTCACAAGG	1860
1801	Db	TCCTCTTAGAGCATCACCTGTACGTAGTCAGTTACGTAAATCTGGAGAGGTCACAAGG	1860
1861	Qy	CTGACTGACCGTGCTACTCACAATTCCTGTGTCATCAGTCAGCATCTCTTTATA	1920
1861	Db	CTGACTGACCGTGCTACTCACAATTCCTGTGTCATCAGTCAGCATCTCTTTATA	1920
1921	Qy	AGTAGATATGTAACCAAGAGNATCCACATGTGTGTCCCTTTACAAGCTATCAAGTCT	1980
1921	Db	AGTAGATATGTAACCAAGAGNATCCACATGTGTGTCCCTTTACAAGCTATCAAGTCT	1980
1981	Qy	GAAGATGCCAACTTGGCAAAACAAGGAATTTTGGGCCACCATTTTGGATTCAGAGGT	2040
1981	Db	GAAGATGCCAACTTGGCAAAACAAGGAATTTTGGGCCACCATTTTGGATTCAGAGGT	2040
2041	Qy	CCTCTTCTGACTATACCTCTCGAATAATTTTCTTTTGAAGTACTACTCTGGATTTACA	2100
2041	Db	CCTCTTCTGACTATACCTCTCGAATAATTTTCTTTTGAAGTACTACTCTGGATTTACA	2100
2101	Qy	TTGTATGGGATGCTCTACAAGCCCTCATGATCTACAGCCCTGGAAAGAAATATCTCTACTGTG	2160
2101	Db	TTGTATGGGATGCTCTACAAGCCCTCATGATCTACAGCCCTGGAAAGAAATATCTCTACTGTG	2160
2161	Qy	CTGTTTCATATATGGTGGTCTCTCAGGTGTCAGTTGGTGAATAATCTTAAAGGAGTCAAG	2220
2161	Db	CTGTTTCATATATGGTGGTCTCTCAGGTGTCAGTTGGTGAATAATCTTAAAGGAGTCAAG	2220
2221	Qy	TATTTCCGCTTGAATACCTTAGCTCTCTAGGTTTATGTGGTTGTAGTAGACAAACAGG	2280
2180	Db	-----	2179
2281	Qy	GGATCTGTACCGAGGGCTTAAATTTTGAAGGGCGCTTTTAAATATAAAATCGGTCAATA	2340

Db	2180	-----GGGTCAATA 2181
Qy	2341	GAATTTGACGATACAGGTGGAGGATCCCAATATCTAGCTTCTCGATATGATTCTCATTCAC 2400
Db	2190	GAATTTGACGATACAGGTGGAGGATCCCAATATCTAGCTTCTCGATATGATTCTCATTCAC 2249
Qy	2401	TTAGATCTGTGGGATCCACCGCTGGTCTATPGAGGATACCTCTCCCTCGATGGCATT 2460
Db	2250	TTAGATCTGTGGGATCCACCGCTGGTCTATPGAGGATACCTCTCCCTCGATGGCATT 2309
Qy	2461	ATGCAGAGGTGAGATATCTTTCAGGGTTGCTATTGCTGGGGCCGCCAGTCACTCTGTGGATC 2520
Db	2310	ATGCAGAGGTGAGATATCTTTCAGGGTTGCTATTGCTGGGGCCGCCAGTCACTCTGTGGATC 2369
Qy	2521	TTCTATGATACAGGATACACGGAACGTTTATATGGGTCACTCTGACACAGAAATGAACAGGGC 2580
Db	2370	TTCTATGATACAGGATACACGGAACGTTTATATGGGTCACTCTGACACAGAAATGAACAGGGC 2429
Qy	2581	TATTACTTAGGATCTGTGGCCATCGAAGAGAAAGTTCCTCTGAAACCAATCGTTT 2640
Db	2430	TATTACTTAGGATCTGTGGCCATCGAAGAGAAAGTTCCTCTGAAACCAATCGTTT 2489
Qy	2641	CTGCTCTTACATGGTTTCTCGATGAGAAATGTCATTTTGGACATACCAAGTATATTCTG 2700
Db	2490	CTGCTCTTACATGGTTTCTCGATGAGAAATGTCATTTTGGACATACCAAGTATATTCTG 2549
Qy	2701	AGTTTTTTAGTGAAGGCTGGAAAGCCATATGATTTACAGATCTATCTCCAGGAGACAC 2760
Db	2550	AGTTTTTTAGTGAAGGCTGGAAAGCCATATGATTTACAGATCTATCTCCAGGAGACAC 2609
Qy	2761	AGCAATAAGATTCCTGAATCGGGAGAACATTTATGAACCTGCATCTTTTGACACTACCTTCA 2820
Db	2610	AGCAATAAGATTCCTGAATCGGGAGAACATTTATGAACCTGCATCTTTTGACACTACCTTCA 2669
Qy	2821	GAACAACCTTGGATCACTGATTCCTGCTCTTAAAAAGTGATATAATTTTGACCTGTGTAGAAC 2880
Db	2670	GAACAACCTTGGATCACTGATTCCTGCTCTTAAAAAGTGATATAATTTTGACCTGTGTAGAAC 2729
Qy	2881	TCTCTGATATACACTGGCTATTAAACCAATGAGGAGGTTTAAATCAACAGAAAAACACAGA 2940
Db	2730	TCTCTGATATACACTGGCTATTAAACCAATGAGGAGGTTTAAATCAACAGAAAAACACAGA 2789
Qy	2941	ATTGATCATCAATTTTGATACCTGCCATGTAAACATCTACTCTCGAAAAATAAATGTGTGTG 3000
Db	2790	ATTGATCATCAATTTTGATACCTGCCATGTAAACATCTACTCTCGAAAAATAAATGTGTGTG 2849
Qy	3001	CCATGACAGGGTCTACGGTTTGTGGTAGTAATCTAATACCTTAAACCCACATGCTCAAAA 3060
Db	2850	CCATGACAGGGTCTACGGTTTGTGGTAGTAATCTAATACCTTAAACCCACATGCTCAAAA 2909
Qy	3061	TCAAATGATACATATCTCTGAGAGACCCAGCAATACCATAGAAATTTACTTAAAAAATAA 3120
Db	2910	TCAAATGATACATATCTCTGAGAGACCCAGCAATACCATAGAAATTTACTTAAAAAATAA 2969
RESULT 11		
AX608743		
LOCUS	4676 bp	DNA linear PAT 17-FEB-2003
DEFINITION	Sequence 20 from Patent WO0231134.	
ACCESSION	AX608743	
VERSION	AX608743.1	GI:28404307
KEYWORDS	Homo sapiens (human)	
SOURCE	Homo sapiens	
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.	
REFERENCE 1	Qi, S., Akinsanya, K.O., Riviere, P.J. and Junien, J.L.	
AUTHORS	Novel serine protease genes related to dppiv	
TITLE	Patent: WO 0231134-A 20 18-APR-2002;	
JOURNAL	Ferring BV (NL)	
FEATURES	Location/Qualifiers	

source	1. .4676																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																				
--------	----------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Qy	2041	CCTCTTCTGACTACTCTCCAGAAATTTCTCTTTTGAAGTACTACTGGAATTTACA	2100
Db	2041	CCTCTTCTGACTACTCTCTCCAGAAATTTCTCTTTTGAAGTACTACTGGAATTTACA	2100
Qy	2101	TTGTATGGGATGCTCTACAGCCTCATGATCTACAGCCTGGAAGAAATATCTTACTGTG	2160
Db	2101	TTGTATGGGATGCTCTACAGCCTCATGATCTACAGCCTGGAAGAAATATCTTACTGTG	2160
Qy	2161	CTGTTTCAATATATGTTGCTCTCAGGTGCAGTTCGGTGAATAATCGGTTTAAAGGAGTCAAG	2220
Db	2161	CTGTTTCAATATATGTTGCTCTCAGGTGCAGTTCGGTGAATAATCGGTTTAAAGGAGTCAAG	2220
Qy	2221	TATTTCCGCTTTGAATACCTAGCCTCTCTAGGTTATGTGGTTGTAGTGATAGACAACAGG	2280
Db	2180	-----	2179
Qy	2281	GGATCCTGTGACCAGGGGCTTAAATTTGAAGGCGCCTTTAAATATATAAATGGGTCAATA	2340
Db	2180	-----GGGTCAATAA	2189
Qy	2341	GAAATGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTCAATTGAC	2400
Db	2190	GAAATGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTCAATTGAC	2249
Qy	2401	TTAGATCTGTGGGCATCCAGCGTGTCTCTATGAGGATACCTCTCCCTGATGCAATTA	2460
Db	2250	TTAGATCTGTGGGCATCCAGCGTGTCTCTATGAGGATACCTCTCCCTGATGCAATTA	2309
Qy	2461	ATGCAGAGGTGAGATATCTTCAGGGTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGATC	2520
Db	2310	ATGCAGAGGTGAGATATCTTCAGGGTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGATC	2369
Qy	2521	TTCTATGATACAGGATACAGGACGCTTATATGGGTGCTATGCTGGGGCCCCAGTCACTCTGTGGATC	2580
Db	2370	TTCTATGATACAGGATACAGGACGCTTATATGGGTGCTATGCTGGGGCCCCAGTCACTCTGTGGATC	2429
Qy	2581	TATTACTTAGATCTGTGGCCATCCAGCAGCAAAAGTTCCCTCTGCAACCAATCGTTTA	2640
Db	2430	TATTACTTAGATCTGTGGCCATCCAGCAGCAAAAGTTCCCTCTGCAACCAATCGTTTA	2489
Qy	2641	CTGCTCTTACATGTTCTCTGGATGAGAAATGTCATTTTGCACATACCAATATATCTG	2700
Db	2490	CTGCTCTTACATGTTCTCTGGATGAGAAATGTCATTTTGCACATACCAATATATCTG	2549
Qy	2701	AGTTTTTTAGTGAGGGCTGGAAGCCCATATGATTTACAGATCTATCTCTCAGGAGAGAC	2760
Db	2550	AGTTTTTTAGTGAGGGCTGGAAGCCCATATGATTTACAGATCTATCTCTCAGGAGAGAC	2609
Qy	2761	AGCATAAGAGTTCTCTGAATCGGGAGAACATTAATGAATGCACTCTTTTGCACACTCTCAA	2820
Db	2610	AGCATAAGAGTTCTCTGAATCGGGAGAACATTAATGAATGCACTCTTTTGCACACTCTCAA	2669
Qy	2821	GAAGACCTTGGATCAGTATTTGCTCTTAAAGTGATATAATTTTGAACCTGTGTAGAAC	2880
Db	2670	GAAGACCTTGGATCAGTATTTGCTCTTAAAGTGATATAATTTTGAACCTGTGTAGAAC	2729
Qy	2881	TCTCTGGTATACACTGGCTATTTAAACCAATGAGGAGTTTAAATCAACGAAACACAGA	2940
Db	2730	TCTCTGGTATACACTGGCTATTTAAACCAATGAGGAGTTTAAATCAACGAAACACAGA	2789
Qy	2941	ATTGATCATCATTTTGCATCTGACCTGATATCTCTCTGAAATTAATGTGGTG	3000
Db	2790	ATTGATCATCATTTTGCATCTGACCTGATATCTCTCTGAAATTAATGTGGTG	2849
Qy	3001	CAATGAGGGCTACGGTTTGTGGTAGTAACTTAATACCTTAACCCACATGCTCAAA	3060
Db	2850	CAATGAGGGCTACGGTTTGTGGTAGTAACTTAATACCTTAACCCACATGCTCAAA	2909
Qy	3061	TCAATATGATACATATCTCTGAGAGACCCAGGATACCATTAAGATTAATCAAAAAA	3120
Db	2910	TCAATATGATACATATCTCTGAGAGACCCAGGATACCATTAAGATTAATCAAAAAA	2969
RESULT 12			
AY354202			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
gene			
CDS			
Location/Qualifiers			
1..3030			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/chromosome="15"			
/tissue_type="testis"			
1..3030			
/gene="Dpp8"			
161..2857			
/gene="Dpp8"			
/note="alternatively spliced"			
/codon_start=1			
/product="dipeptidyl peptidase 8 isoform 3"			
/protein_id="AAQ63887.1"			
/db_xref="GI:34329352"			
/translation="MKRSKQADTKKVKYGMMAKAPHDPMFVKRNDPDGPHSDRIYVL			
RPKLPFVRYVSWSLKDLKADTKKVKYGMMAKAPHDPMFVKRNDPDGPHSDRIYVL			
AMSGNRNTLPYSEIPIKTINRAAVLMSKPLLDLPQATLDYGMVSRBELLRKR			
TGTGIVASVDYHQSGTFLQAGSGIVHVKGQPGQPTQPLRPNVETSCPNRMDP			
KLCPADPWIAFTHSNDIMSNIVTREERLTIVHNELEEDASAGVATFVLOEE			
FDRYGVWVKPKARTTSGKILALILEENDESEVELIHTVSPLMTRRADSPRYPKT			
GTANPKVTFKSGEIMIDAEGLIDVIOKELIOPPELLFEGEVIARAGWTPEGKYAMS			
ILLDRSQRLQIVLISLELTFPVEDDMORORLIESVPDSVTPLIYEETDIIWINIH			
DIFHVPDSHBEIEFIFASECKTGFRHLKXTSLIKESKYKSSGGLVAPSPKCP			
KEEIAITSGEWEVIGHGNSIQVDEVRRLVYFEGTKDSPLIEHLKYVSVYVNPGEVTEL			
TDGYSHSCCSIQHCDPFISKYSNQKPHCVSLYKQSSPDEDDTCKTKERWATILISA			
GLPDYTPPEIFPSEFTTFTLYGLKPHDLQPKKIFIVLFIYGGFQQLVNNRKF			
GVKYFRLNTLASLGVVVVINDRGSCHRLKPEGAFKMGQIEIDDOVREGQLIASR			
YDFIDLRVGIHGWYGGYLSLMALMQRSDFRVAIAGAPVILWIFIDTGTIERIGH			
PDNSQGYLGSVAQAEKFPSENRLLLLHGLDENVHFATPILLSLLFLVRAGKPYD			
LQIYPOERHSIRVPESGHEYLHLLHLYQENLGSRIAALKVI"			
ORIGIN			
Query Match 89.9%; Score 2805.4; DB 8; Length 3030;			
Best Local Similarity 99.9%; Pred. No. 0;			
Matches 2806; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
Qy	201	AGGAAATGCAACATGGCAGCAGCAATGGAAACAGAACAGCTGGGTGTTGAGATATTGA	260
Db	196	AGGAAATGCAACATGGCAGCAGCAATGGAAACAGAACAGCTGGGTGTTGAGATATTGA	255
Qy	261	AACCTGGGACTGTGAGGAGAAATTAATTAATCAAGATCGGCTAAATGGAGCCTTTTA	320
Db	256	AACCTGGGACTGTGAGGAGAAATTAATTAATCAAGATCGGCTAAATGGAGCCTTTTA	315
Qy	321	TGTTGAGCGGTATCTCTGAGTCAGCTTAAAGAGCTGCTGCCATACCCAGAAATATCA	380
Db	316	TGTTGAGCGGTATCTCTGAGTCAGCTTAAAGAGCTGCTGCCATACCCAGAAATATCA	375
Qy	381	TGCTACATGATGGCTTAAGGCACCATGATTTTCTGTAAGAGGAGGATGATCCAGA	440
Db	376	TGCTACATGATGGCTTAAGGCACCATGATTTTCTGTAAGAGGAGGATGATCCAGA	435

Qy 441 TGAACCTCATTCAGACAGAACTCTATTACCTTGCCATGTCTGGTGAGAACAGAGAAATATAC 500
Db 436 TGAACCTCATTCAGACAGAACTCTATTACCTTGCCATGTCTGGTGAGAACAGAGAAATATAC 495
Qy 501 ACTGTTTATCTGAAATTTCCCAAACTATCAATAGAGCAGCAGTCTTAAATGCTCTCTTG 560
Db 496 ACTGTTTATCTGAAATTTCCCAAACTATCAATAGAGCAGCAGTCTTAAATGCTCTCTTG 555
Qy 561 GAAGCCTCTTTTGGATCTTTTTCAGGCAACACTGGACTATGGAAATGATTTCTCGAGAAGA 620
Db 556 GAAGCCTCTTTTGGATCTTTTTCAGGCAACACTGGACTATGGAAATGATTTCTCGAGAAGA 615
Qy 621 AGAATATTAAAGAAAGAAACGCAATTTGGAACAGTGGAAATTCCTTACCAATATCA 580
Db 616 AGAATATTAAAGAAAGAAACGCAATTTGGAACAGTGGAAATTCCTTACCAATATCA 675
Qy 681 CCAAGCAAGTGGAAATTTCTGTTTCAAGCCGGTAGTGGAAATTTATCACGTAAAGATGG 740
Db 676 CCAAGCAAGTGGAAATTTCTGTTTCAAGCCGGTAGTGGAAATTTATCACGTAAAGATGG 735
Qy 741 AGGGCCCAAGGATTTACGCAACAACTTTTAAGGCCCAATCTAGTGGAAACTAGTTGTCC 800
Db 736 AGGGCCCAAGGATTTACGCAACAACTTTTAAGGCCCAATCTAGTGGAAACTAGTTGTCC 795
Qy 801 CAAATACCGATGGATCCAAAATTAATGCCCGCTGATCCAGACTGGATTCCTTTATACA 860
Db 796 CAAATACCGATGGATCCAAAATTAATGCCCGCTGATCCAGACTGGATTCCTTTATACA 855
Qy 861 TAGCAACGATATTTGGATATCTAAATCGTAAACAGAGAAAGGAGACTCACTTATGT 920
Db 856 TAGCAACGATATTTGGATATCTAAATCGTAAACAGAGAAAGGAGACTCACTTATGT 915
Qy 921 GCACAATGAGCTAGCCCAATGGAAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTGT 980
Db 916 GCACAATGAGCTAGCCCAATGGAAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTGT 975
Qy 981 TCTCCAAAGAAATTTGATAGATATTTCTGGCTATTTGGTGTGTCCTCAAAAGCTGAACAAC 1040
Db 976 TCTCCAAAGAAATTTGATAGATATTTCTGGCTATTTGGTGTGTCCTCAAAAGCTGAACAAC 1035
Qy 1041 TCCAGTGGTGAATTTCTAGAAATCTATATGAAGAAATGATGAATCTGAGGTGA 1100
Db 1036 TCCAGTGGTGAATTTCTAGAAATCTATATGAAGAAATGATGAATCTGAGGTGA 1095
Qy 1101 AATTATCATGTTACATCCCTTATGTTGMAACAGGAGGCGAGATTCATTCCGTTATCC 1160
Db 1096 AATTATCATGTTACATCCCTTATGTTGMAACAGGAGGCGAGATTCATTCCGTTATCC 1155
Qy 1161 TAAACAGGTACAGCAAACTTAAAGTCACTTTTAAAGATGTCAGAAATAATGATGATGC 1220
Db 1156 TAAACAGGTACAGCAAACTTAAAGTCACTTTTAAAGATGTCAGAAATAATGATGATGC 1215
Qy 1221 TGAAGGAAGATCATAGATGTATAGATAAGAACTAATTCACCTTTTGAGATTCATTT 1280
Db 1216 TGAAGGAAGATCATAGATGTATAGATAAGAACTAATTCACCTTTTGAGATTCATTT 1275
Qy 1281 TGAAGAGTTGAATATATTGCCAGAGCTGGATGGACTCTCAGGGAATAATGCTGGTC 1340
Db 1276 TGAAGAGTTGAATATATTGCCAGAGCTGGATGGACTCTCAGGGAATAATGCTGGTC 1335
Qy 1341 CATCCTACTAGATCGCTCCAGACTCGCCTACAGATAGTGTGATCTCACTCGAATATT 1400
Db 1336 CATCCTACTAGATCGCTCCAGACTCGCCTACAGATAGTGTGATCTCACTCGAATATT 1395
Qy 1401 TATCCAGTAGAAGATGATGTTTATGGAAGGAGAGACTCATTCAGTCACTGCTGATTC 1460
Db 1396 TATCCAGTAGAAGATGATGTTTATGGAAGGAGAGACTCATTCAGTCACTGCTGATTC 1455
Qy 1461 TGTGAGGCCCACTAATTTATCTATGAAGAAACAAACAGACATCTGGATAAATATCCATGACAT 1520
Db 1456 TGTGAGGCCCACTAATTTATCTATGAAGAAACAAACAGACATCTGGATAAATATCCATGACAT 1515
Qy 1521 CTTTCATGTTTTTCCCAAAAGTCACGAAGAGAAATTCAGTTTATTTTTTGCTCTGAATG 1580

Db 1516 CTTTCATGTTTTTCCCAAAAGTCACGAAGAGAAATTCAGTTTATTTTTTGCTCTGAATG 1575
Qy 1581 CAAAAACAGGTTTCCGTCATTATACAAAATTCATCTATTTTAAAGGAAAGCAAAATATAA 1640
Db 1576 CAAAAACAGGTTTCCGTCATTATACAAAATTCATCTATTTTAAAGGAAAGCAAAATATAA 1635
Qy 1641 ACATCCAGTGGTGGGCTGCCCTGCTCCAAAGTGATTTCAAGTGTCTCTATCAAGAGGAGAT 1700
Db 1636 ACATCCAGTGGTGGGCTGCCCTGCTCCAAAGTGATTTCAAGTGTCTCTATCAAGAGGAGAT 1695
Qy 1701 AGCAATTTACAGTGGTGAATGGAAGTTCCTGGCCGCAATGGATCTAATATATCAAGTTGA 1760
Db 1696 AGCAATTTACAGTGGTGAATGGAAGTTCCTGGCCGCAATGGATCTAATATATCAAGTTGA 1755
Qy 1761 TGAAGTCAGAGGCTGGTATATTTTTGAAGGCCAACAAAGACTCCCTTTTGAAGCATCACCT 1820
Db 1756 TGAAGTCAGAGGCTGGTATATTTTTGAAGGCCAACAAAGACTCCCTTTTGAAGCATCACCT 1815
Qy 1821 GTACGTAGTCAGTTACGTAAATCTCGAGAGGTGAACAAGGTGACTGACCGTGGCTACTC 1880
Db 1816 GTACGTAGTCAGTTACGTAAATCTCGAGAGGTGAACAAGGTGACTGACCGTGGCTACTC 1875
Qy 1881 ACATTTCTGTCGATCAGTCAGACCTGTGACTCTTTTATAGTAAGTATAGTAACCGAGAA 1940
Db 1876 ACATTTCTGTCGATCAGTCAGACCTGTGACTCTTTTATAGTAAGTATAGTAACCGAGAA 1935
Qy 1941 GAATCCACACTGTGTCCCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCCACTTGCAA 2000
Db 1936 GAATCCACACTGTGTCCCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCCACTTGCAA 1995
Qy 2001 AACAAAGGAATTTTGGGCCACCAATTTTGGATTCAGAGGTCTCTTCTCTGACTATCTCC 2060
Db 1996 AACAAAGGAATTTTGGGCCACCAATTTTGGATTCAGAGGTCTCTTCTCTGACTATCTCC 2055
Qy 2061 TCCAGAAATTTTCTCTTTTGAAGTACTACTGATTTACATTTGATGGGATGCTCTACAA 2120
Db 2056 TCCAGAAATTTTCTCTTTTGAAGTACTACTGATTTACATTTGATGGGATGCTCTACAA 2115
Qy 2121 GCCTCATGATCTACGCTGGAAAGAAATATCTACTGTCTGTGTTTATATATGTTGGTGC 2180
Db 2116 GCCTCATGATCTACGCTGGAAAGAAATATCTACTGTCTGTGTTTATATATGTTGGTGC 2175
Qy 2181 TCAGGTGCACTGGTGAATAATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCCCT 2240
Db 2176 TCAGGTGCACTGGTGAATAATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCCCT 2235
Qy 2241 AGCCTCTTAGGTTATGTGGTTGATGATAGACAAACAGGGGATCCTGTCAACCGAGGCT 2300
Db 2236 AGCCTCTTAGGTTATGTGGTTGATGATAGACAAACAGGGGATCCTGTCAACCGAGGCT 2295
Qy 2301 TAAATTTGAAGGCGCTTTTAAATATATAAATGGTCAAAATAGAAATTTGACGATCAGGTGGA 2360
Db 2296 TAAATTTGAAGGCGCTTTTAAATATATAAATGGTCAAAATAGAAATTTGACGATCAGGTGGA 2355
Qy 2361 AGGACTCCAATATCTAGCTTTCTCGATATGATTTCAATTTAGACTTTAGATCGTGGGATCCA 2420
Db 2356 AGGACTCCAATATCTAGCTTTCTCGATATGATTTCAATTTAGACTTTAGATCGTGGGATCCA 2415
Qy 2421 CGGCTGTCTTATGGAAGGATACCTCTCCCTGATGGCAATTAATGCAAGAGTCAATATCTTT 2480
Db 2416 CGGCTGTCTTATGGAAGGATACCTCTCCCTGATGGCAATTAATGCAAGAGTCAATATCTTT 2475
Qy 2481 CAGGTTGCTATTTGCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACAC 2540
Db 2476 CAGGTTGCTATTTGCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACAC 2535
Qy 2541 GGAACGTTATATGGGTACCCCTGACCAAGATGAACAGGGCTATTTACTTAGGATCTGTGGC 2600
Db 2536 GGAACGTTATATGGGTACCCCTGACCAAGATGAACAGGGCTATTTACTTAGGATCTGTGGC 2595
Qy 2601 CATCAAGCAGAAAGTTCCCTCTCGAAACAAATCGTTTACTGTCTTCTACATGTTTCTCT 2660

Db 1381 CAGTGGTGAATGGGAAGTTCTTGGCCGCATGGATCTAATATCCAAAGTTGATGAAGTCAG 1440
Qy 1770 AAGGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCACTGTACGTAGT 1829
Db 1441 AAGGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCACTGTACGTAGT 1500
Qy 1830 CAGTTACGTAAATCTCGAGAGGTGACAAAGCTGACCTGACCGTGTACTCATTCTTTG 1889
Db 1501 CAGTTACGTAAATCTCGAGAGGTGACAAAGCTGACCTGACCGTGTACTCATTCTTTG 1560
Qy 1890 CTGCATCAGTCAGCACTGTGACTCTCTTTATAAGTAAGTATAGTAACCAAGAAATCCACA 1949
Db 1561 CTGCATCAGTCAGCACTGTGACTCTCTTTATAAGTAAGTATAGTAACCAAGAAATCCACA 1620
Qy 1950 CTGTGTGTCCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAGAAAGGA 2009
Db 1621 CTGTGTGTCCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAGAAAGGA 1680
Qy 2010 ATTTTGGGCGACCATTTTGGATTTGAGATTCAGAGTCTCTTCTGATATCTCTCCAGAAAT 2069
Db 1681 ATTTTGGGCGACCATTTTGGATTTGAGATTTGAGATTCAGAGTCTCTTCTGATATCTCTCCAGAAAT 1740
Qy 2070 TTTCTCTTTTGAAGTACTCTGATTTTACATTTGATGGATGCTCTACAAGCCTCATGA 2129
Db 1741 TTTCTCTTTTGAAGTACTCTGATTTTACATTTGATGGATGCTCTACAAGCCTCATGA 1800
Qy 2130 TCTACAGCTCGAAGAAATATCTCTGCTGCTGCTCATATATGATGGTCTCTCAGGTGCA 2189
Db 1801 TCTACAGCTCGAAGAAATATCTCTGCTGCTGCTCATATATGATGGTCTCTCAGGTGCA 1860
Qy 2190 GTTGGTGAATATCGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTCAAGCTCTCT 2249
Db 1861 GTTGGTGAATATCGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTCAAGCTCTCT 1920
Qy 2250 AGGTTATGTTGGTGTAGTGATAGACAAACAGGGGATCTGTCAACGAGGGCTTAAATTTGA 2309
Db 1921 AGGTTATGTTGGTGTAGTGATAGACAAACAGGGGATCTGTCAACGAGGGCTTAAATTTGA 1980
Qy 2310 AGGCGCTTTAAATATAAAATGGGTCAAAATAGAAATTTGACATCAGGTGGAAGGACTCCA 2369
Db 1981 AGGCGCTTTAAATATAAAATGGGTCAAAATAGAAATTTGACATCAGGTGGAAGGACTCCA 2040
Qy 2370 ATATCTAGCTTCTCGATATGATTTCAATTTGATAGATCGTGTGGGCATCCAGGCTGGTC 2429
Db 2041 ATATCTAGCTTCTCGATATGATTTCAATTTGATAGATCGTGTGGGCATCCAGGCTGGTC 2100
Qy 2430 CTATGAGGATACCTCTCCCTGATGGCAATTAATGACAGGTCTAGATATCTTCAGGGTGGC 2489
Db 2101 CTATGAGGATACCTCTCCCTGATGGCAATTAATGACAGGTCTAGATATCTTCAGGGTGGC 2160
Qy 2490 TATTGCTGGGGCCCGAGTCACCTCTGTGTGATCTCTATGATACAGGATACACGGAACGTGA 2549
Db 2161 TATTGCTGGGGCCCGAGTCACCTCTGTGTGATCTCTATGATACAGGATACACGGAACGTGA 2220
Qy 2550 TATGGGTACCTCTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGCCATGCAAGC 2609
Db 2221 TATGGGTACCTCTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGCCATGCAAGC 2280
Qy 2610 AGAAAAGTTCCCTCTGAAACCAATCGTTTACTGCTTTACATGTTTCTCGATGAGAA 2669
Db 2281 AGAAAAGTTCCCTCTGAAACCAATCGTTTACTGCTTTACATGTTTCTCGATGAGAA 2340
Qy 2670 TGTCCATTTTGCACATACCAAGTATTAATCTGAGTTTTTATGAGGCTGGAAAGCCATA 2729
Db 2341 TGTCCATTTTGCACATACCAAGTATTAATCTGAGTTTTTATGAGGCTGGAAAGCCATA 2400
Qy 2730 TGATTTACAGATCTATCTCTCAGGAGAGACACAGCATATAGAGTCTCTGATCGGGAGAAC 2789
Db 2401 TGATTTACAGATCTATCTCTCAGGAGAGACACAGCATATAGAGTCTCTGATCGGGAGAAC 2460
Qy 2790 TTATGAATCGCATCTTTTGGCACTACTCTCAGAAAAACCTTGGATCAGTATTCCTGCTCT 2849
Db 2461 TTATGAATCGCATCTTTTGGCACTACTCTCAGAAAAACCTTGGATCAGTATTCCTGCTCT 2520

Qy 2850 AAAAGTGATATAATTTTGACCTGTGTAGAACTCTCTGTGTATACACTGGCTATTTAACCAA 2909
Db 2521 AAAAGTGATATAATTTTGACCTGTGTAGAACTCTCTGTGTATACACTGGCTATTTAACCAA 2580
Qy 2910 ATGAGGAGGTTTAAATCAACAGAAAAACACAGAAATGATCATCATTTTGTGATCTGCGAT 2969
Db 2581 ATGAGGAGGTTTAAATCAACAGAAAAACACAGAAATGATCATCATTTTGTGATCTGCGAT 2640
Qy 2970 GTPAACATCTACTCTCGAAAAATAAATCTGTGTCATGAGGGGTCTACGGTTTGTGGTAGT 3029
Db 2641 GTAACATCTACTCTCGAAAAATAAATCTGTGTCATGAGGGGTCTACGGTTTGTGGTAGT 2700
Qy 3030 AATCTAATACCTTAACCCACATGCTCAAAATCAAAATGATACATATTTCTCTGAGAGCCCA 3089
Db 2701 AATCTAATACCTTAACCCACATGCTCAAAATCAAAATGATACATATTTCTCTGAGAGCCCA 2760
Qy 3090 GCAATACCAATAAGAAATTACTTAAAAAATAA 3120
Db 2761 GCAATACCAATAAGAAATTACTTAAAAAATAA 2791

RESULT 14

BC030688 3125 bp mRNA linear PRI 28-JUL-2005
Homo sapiens dipeptidylpeptidase 8, transcript variant 4, mRNA
(cDNA clone MGC:26191 IMAGE:4822550), complete cds.

ACCESSION BC030688.2 GI:34190028
VERSION BC030688.2
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE 1 (bases 1 to 3125)

Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Haieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Sapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Schetz, T.E., Brownstein, M.J., Usdin, T.B., Toohy, S.,
Carninci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahney, J., Helton, E., Kettner, M., Madan, A., Rodrigues, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shvachenko, Y.,
Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalek, U., Smal, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.

Mammalian Gene Collection Program Team

Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

2 (bases 1 to 3125)

NIH MGC Project

Direct Submission

Submitted (24-MAY-2002) National Institutes of Health, Mammalian

Gene Collection (MGC), Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

On Aug '05, 2003 this sequence version replaced gi:21265132.

Contact: MGC help desk

Email: cgaps-remail.nih.gov

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.

cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki

Toshiyuki and Piero Carninci (RIKEN)

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Institute for Systems Biology
http://www.systemsbio.org
contact: amadan@systemsbio.org
Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha Madan, Stephanie Rodriguez, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 33 Row: d Column: 5
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 37577092.

FEATURES

Location/Qualifiers
1..3125

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clones="MGC:26191 IMAGE:4822550"
/tissue_type="Testis"
/clone_lib="NIH MGC_97"
/lab_host="DH10B"
/note="vector: pBluescriptR"

gene

1..3125
/gene="DPP8"
/name="synonyms: FLJ20283, FLJ14920, MGC26191, MSTP141"
/db_xref="GeneID:54878"
319..2862

CDS

/gene="DPP8"
/codon_start=1
/product="dipeptidyl peptidase 8, isoform 4"
/protein_id="AAH30688.3"
/db_xref="GI:54114905"
/db_xref="GeneID:54878"
/db_xref="MIM:606819"
/translation="MWKRSQMKIKSGKCNMAAMETQLGVEIPETADCBENIESQD
RPKLEPFVVERYSQKLLADTRKYHGYMMAKPHDFVFRKNDPDGPHSDRIYYL
AMSGENRENTLFYSEI PKTINRAAIVMLSKPLDLFQATLDYGMYSREBELRERKR
IGTGVIASYDHQSGTFTFAAGSGYHVKGQPGQFTQOPLRPNLVSTSPNIRMDP
KLCPADPWIAFHSNDIWI SNITVREERLYVHNEELMEEDARSAGVATFVLQEE
FDYSGYVMCKPASTTPSGKKILRIYENDESEVEI IHTVSPMLSTRADSFRYPKT
GTAMPKVTKMSIEMDAEGRIIDVIDKELIQPEILFEGVEYIARAGWTPGKYAWS
ILLDRSQRILQVILISPELFIPEVDDVMERQLIESVPDSVPLIIEYETDIIWNIH
DIHFVFGSHEEETEFIFASECKTGFRHLVYVPGTKDSPLEHHLVSVVNPGEVTRL
KEEIALTSGSEVLGRHSGNIQVDEVRLLVYVPGTKDSPLEHHLVSVVNPGEVTRL
TDGVSHSCCISOHCDFFI SKYNKNPHCVSLYKLSPPEDDPTCKTKEFWATILDSA
GPLDYPPEI FSPSTTGTLLGMLYKPHDLPQKKYPTLVFIYGGQVOLVNNRPF
GVKVRNLTLASLGYYVVDNRSGCHGLKFGAPKVRVAIAGAPVTLMTIFVDTGY
TERYKHPDQNEQGYLGSVAMQAEKFPSEPNRLILHGLFDENVHFATSIILLSFLV
RAGKPYDLQITPQBRHSIRVPSESEHYELHLLHYLQENLGSRIALKVYI"

ORIGIN

Query Match 85.1%; Score 2656.4; DB 8; Length 3125;
Best Local Similarity 91.1%; Pred. No. 0;
Matches 2966; Conservative 0; Mismatches 1; Indels 289; Gaps 2;
1 AAGTGCTAAAGCCTCCGAGGCCAAGCGCGCTGCTACTGCGCGCGCTGCTTCTTAGTGGCG 60
18 AAGTGCTAAAGCCTCCGAGGCCAAGCGCGCTGCTACTGCGCGCGCGCTGCTTCTTAGTGGCG 77
61 CGTTCCGCGCCTGGGTTTGTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
78 CGTTCCGCGCCTGGGTTTGTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 137
121 GAGTGAGCG 180
138 GAGTGAGCG 197
181 TCCGGCG 200
198 TCCGGCG 257
201 ----- 200

258 ASCAGGATGACGACGAGGGATTCTTATGCTTGAAGTCGAGTCACCTTGAAGAAAGATCTCTTTG 317
201 -----AGGAAATGCAACATGCGCAGCAGC 224
318 ANTGTGGAAGAGATCTGACGAGATGAAATAAATAATCAGGAAATGCAACATGCGCAGCAGC 377
225 AATGAAACAGAACAGCTGGGTGTGAGATATTTTGAACCTGCGGACTGTGTGAGGAGAAATAT 284
378 AATGAAACAGAACAGCTGGGTGTGAGATATTTTGAACCTGCGGACTGTGTGAGGAGAAATAT 437
285 TGAATCAGAGATCGGCTAAATTTGAGAGCCCTTTTATGTTGAGCGGTATTCCTGGAGTCA 344
438 TGAATCAGAGATCGGCTAAATTTGAGAGCCCTTTTATGTTGAGCGGTATTCCTGGAGTCA 497
345 GCTTAAAGAGCTGCTTCCGATACAGAAATAATCATGGCTACATGATGCTAAGGACACC 404
498 GCTTAAAGAGCTGCTTCCGATACAGAAATAATCATGGCTACATGATGCTAAGGACACC 557
405 ACATGATTTTCATGTTTGTGAAAGAGGAATGATCCAGATGGACCTCATTCAGACAGAAATCTA 464
558 ACATGATTTTCATGTTTGTGAAAGAGGAATGATCCAGATGGACCTCATTCAGACAGAAATCTA 617
465 TTACCTTGCCATGCTCGGTGAGAAACAGAGAAATAATACACTGTTTATTTCTGAAATTTCCCAA 524
618 TTACCTTGCCATGCTCGGTGAGAAACAGAGAAATAATACACTGTTTATTTCTGAAATTTCCCAA 677
525 AACTATCAATAGACGACGAGCTTTAATGCTCTCTTGAAGCCCTTTTGGATCTTTTCA 584
678 AACTATCAATAGACGACGAGCTTTAATGCTCTCTTGAAGCCCTTTTGGATCTTTTCA 737
585 GGCACACTGGAATGGAATCTTCCGAGAGAGAACTATTAAAGAGAAAGAAACG 644
738 GGCACACTGGAATGGAATCTTCCGAGAGAGAACTATTAAAGAGAAAGAAACG 797
645 CATTTGAAACAGTCGGAATGCTTCTAGGATATACCAAGAAAGTGGAACTTCTGTT 704
798 CATTTGAAACAGTCGGAATGCTTCTAGGATATACCAAGAAAGTGGAACTTCTGTT 857
705 TCAAGCCGTAGTGGAAATTTATACGTAAAGATGAGAGGCGCAACAGGATTTACGCAACA 764
858 TCAAGCCGTAGTGGAAATTTATACGTAAAGATGAGAGGCGCAACAGGATTTACGCAACA 917
765 ACCTTTAAGGCCCAATCTAGTGGAACTAGTTGTTCCCAACATACGGATGATCCAAAT 824
918 ACCTTTAAGGCCCAATCTAGTGGAACTAGTTGTTCCCAACATACGGATGATCCAAAT 977
825 ATGCCCGCTGATCCAGACTGGATGCTTTTATACATAGCAACGATATTTGGATATCTAA 884
978 ATGCCCGCTGATCCAGACTGGATGCTTTTATACATAGCAACGATATTTGGATATCTAA 1037
885 CATCGTAAACAGAGAAAGGAGACTCCTTATGTGCACAAATGAGCTAGCACAACATGGA 944
1038 CATCGTAAACAGAGAAAGGAGACTCCTTATGTGCACAAATGAGCTAGCACAACATGGA 1097
945 AGAGATGCGCAGTCAGTGGGCTGCTACCTTTGTTCTCCCAAGAAATTTGATAGATA 1004
1098 AGAGATGCGCAGTCAGTGGGCTGCTACCTTTGTTCTCCCAAGAAATTTGATAGATA 1157
1005 TTCTGCTATTTGGTGTGTCCTAAAGCTGAAACAACTCCAGTGGTGGTAAATTTCTTAG 1064
1158 TTCTGCTATTTGGTGTGTCCTAAAGCTGAAACAACTCCAGTGGTGGTAAATTTCTTAG 1217
1065 AATTTATATGAAGAAATGATGAATCTGAGGTGAAATTTATTTATGTTTACATCCCTAT 1124
1218 AATTTATATGAAGAAATGATGAATCTGAGGTGAAATTTATTTATGTTTACATCCCTAT 1277
1125 GTTGGAAACAGAGGCGGCGGATTCATTCGTTATCTTAAAAACAGGTACAGCAAAATCTTAA 1184
1278 GTTGGAAACAGAGGCGGCGGATTCATTCGTTATCTTAAAAACAGGTACAGCAAAATCTTAA 1337
1185 AGTCACCTTTTAAGATGTCAGAAATTAATGATTGATGCTGAAGAGAGGATCATAGATGTCAT 1244
1338 AGTCACCTTTTAAGATGTCAGAAATTAATGATTGATGCTGAAGAGAGGATCATAGATGTCAT 1397

QY	1245	AGATAAGGAACAACTTTTCAAGAGGATCTATTTTGAAGAGGTTGAATATATATGCGCAG	1304
Db	1398	AGATAAGGAACAACTTTTCAAGAGGATCTATTTTGAAGAGGTTGAATATATATGCGCAG	1457
QY	1305	AGCTGGATGACCTCTGAGGGAATAATATCTTGGTCCATCTCTAGATCGCTCCACAGAC	1364
Db	1458	AGCTGGATGACCTCTGAGGGAATAATATCTTGGTCCATCTCTAGATCGCTCCACAGAC	1517
QY	1365	TCGCCTACAGATAGTGTGATCTCACTGAAATATTTATCCAGTAGAAGATGATGTTAT	1424
Db	1518	TCGCCTACAGATAGTGTGATCTCACTGAAATATTTATCCAGTAGAAGATGATGTTAT	1577
QY	1425	GGAAAGCAGAGACTCAATTTAGAGTCACTGCTGATCTCTGAGCGCACTAATATATCTATGA	1484
Db	1578	GGAAAGCAGAGACTCAATTTAGAGTCACTGCTGATCTCTGAGCGCACTAATATATCTATGA	1637
QY	1485	AGAAACAAACAGACATCTGGATTAATAATCCATGACATCTTTTATGTTTTCCTCCCAAGTCA	1544
Db	1638	AGAAACAAACAGACATCTGGATTAATAATCCATGACATCTTTTATGTTTTCCTCCCAAGTCA	1697
QY	1545	CGAAGAGGAATTTGAGTTTATTTTGGCTCTCAATGCAAAACAGGTTTTCCTCAATTTATA	1604
Db	1698	CGAAGAGGAATTTGAGTTTATTTTGGCTCTCAATGCAAAACAGGTTTTCCTCAATTTATA	1757
QY	1605	CAAAATTTACATCTATTTTAAAGGAAGCAAAATATAAAGATCCAGTGGTGGGCTGCCTGC	1664
Db	1758	CAAAATTTACATCTATTTTAAAGGAAGCAAAATATAAAGATCCAGTGGTGGGCTGCCTGC	1817
QY	1665	TCCAAGTGATTTCAAGTGCTCTATCAAGAGGAGATAGCAATACCAGTGGTGAATGGGA	1724
Db	1818	TCCAAGTGATTTCAAGTGCTCTATCAAGAGGAGATAGCAATACCAGTGGTGAATGGGA	1877
QY	1725	AGTTCTTGGCGGATGGATCTAATAATCCAAGTGATGAAGTCAAGAGGCTGGTATATTT	1784
Db	1878	AGTTCTTGGCGGATGGATCTAATAATCCAAGTGATGAAGTCAAGAGGCTGGTATATTT	1937
QY	1785	TGAAGGACCAAGAGATCCCTTTAGAGCATCACCTGTAGCTAGTACGTAGTAAATCC	1844
Db	1938	TGAAGGACCAAGAGATCCCTTTAGAGCATCACCTGTAGCTAGTACGTAGTAAATCC	1997
QY	1845	TGGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACATTTCTGCTCATCAGTCAGCA	1904
Db	1998	TGGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACATTTCTGCTCATCAGTCAGCA	2057
QY	1905	CTGTGACTCTTTTATAGTAAAGTATAGTAACAGAGAAATCCACACTGTGTGTCCTTTA	1964
Db	2058	CTGTGACTCTTTTATAGTAAAGTATAGTAACAGAGAAATCCACACTGTGTGTCCTTTA	2117
QY	1965	CAAGCTATCAAGTCTGAAGATGACCAACTTGCAGAAACAAAGGAATTTTGGGCGACCAT	2024
Db	2118	CAAGCTATCAAGTCTGAAGATGACCAACTTGCAGAAACAAAGGAATTTTGGGCGACCAT	2177
QY	2025	TTTGGATTCAGCAGGCTCTCTCTGACTATATCTCTCCAGAAATTTTCTCTTTGAAAG	2084
Db	2178	TTTGGATTCAGCAGGCTCTCTCTGACTATATCTCTCCAGAAATTTTCTCTTTGAAAG	2237
QY	2085	TACTACTGGATTTTACATTTGATGGGATGCTCTCAAGCCTCATGATCTACAGCCTGGAAA	2144
Db	2238	TACTACTGGATTTTACATTTGATGGGATGCTCTACAGCCTCATGATCTACAGCCTGGAAA	2297
QY	2145	GAAATATCTACTGTGCTGTTTCAATATATGGTGGTCTCTCAGGTGAGTGGTGAATATCG	2204
Db	2298	GAAATATCTACTGTGCTGTTTCAATATATGGTGGTCTCTCAGGTGAGTGGTGAATATCG	2357
QY	2205	GTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTAGCCTCTCTAGGTATGTTGTTGT	2264
Db	2358	GTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTAGCCTCTCTAGGTATGTTGTTGT	2417
QY	2265	AGTGATGACAAACAGGGGATCTCTGTCAACCGAGGGCTTAAATTTGAAGGCGCTTTAAATA	2324
Db	2418	AGTGATGACAAACAGGGGATCTCTGTCAACCGAGGGCTTAAATTTGAAGGCGCTTTAAATA	2477

RESULT 15
AX354795
LOCUS AX354795 2649 bp DNA linear PAT 06-FEB-2002
DEFINITION Sequence 3 from Patent WO0179473.
ACCESSION AX354795
VERSION AX354795.1 GI:18619528
KEYWORDS Homo sapiens (human)
SOURCE ORGANISM, Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1
REFERENCE
AUTHORS Meyers, R.A. and Williamson, M.
TITLE 21953, a human prolyl oligopeptidase family member and uses thereof
JOURNAL Patent: WO 0179473-A 3 25-OCT-2001;

Tue Apr 18 08:18:37 2006

FEATURES		Millennium Pharmaceuticals, Inc. (US)	
Source	Location/Qualifiers		
	1..2649		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
ORIGIN			
Query Match		84.88;	Score 2645.8; DB 6; Length 2649;
Best Local Similarity		99.98;	Pred. No. 0;
Matches 2647;		Conservative 0;	Mismatches 2; Indels 0; Gaps 0;
Qy	214	ATGGCAGCAGCATGGAAACAGAACAGCTGGGTGTGATATTTGAAACTGGCGACTGT	273
Db	1	ATGGCAGCAGCATGGAAACAGAACAGCTGGGTGTGATATTTGAAACTGGCGACTGT	60
Qy	274	GAGGAGAAATATGGAATCAGAGATCGGCCCTAAATTTGGAGCCCTTTATTTGAGCGGTAT	333
Db	61	GAGGAGAAATATGGAATCAGAGATCGGCCCTAAATTTGGAGCCCTTTATTTGAGCGGTAT	120
Qy	334	TCCTGGAGTCAGCTTAAAGAGCTGCTTGGCGATACAGAAATATCATGGCTACATGATG	393
Db	121	TCCTGGAGTCAGCTTAAAGAGCTGCTTGGCGATACAGAAATATCATGGCTACATGATG	180
Qy	394	GCTAAGGACCAACATGATTTTCATGTTTGTGAAGAGAAATGATCCAGATGGACCTCATTC	453
Db	181	GCTAAGGACCAACATGATTTTCATGTTTGTGAAGAGAAATGATCCAGATGGACCTCATTC	240
Qy	454	GACAGAACTCTATTACCTTTGCCATGCTCTGGTGAGAACAGAGAAATACTGTTTATTTCT	513
Db	241	GACAGAACTCTATTACCTTTGCCATGCTCTGGTGAGAACAGAGAAATACTGTTTATTTCT	300
Qy	514	GAATTTCCCAAACTATCAATAGAGCAGCAGTCTTAATGCTCTCTTTGGAAGCTCTTTTG	573
Db	301	GAATTTCCCAAACTATCAATAGAGCAGCAGTCTTAATGCTCTCTTTGGAAGCTCTTTTG	360
Qy	574	GATCTTTTCAGGCAACTCGACTATGGAATGATTTCTCGAAGAGAACTATTTAAGA	633
Db	361	GATCTTTTCAGGCAACTCGACTATGGAATGATTTCTCGAAGAGAACTATTTAAGA	420
Qy	634	GAAGAAACCGATTTGGACACAGTCGGAAATGCTTCTTACGATATACCAAGAGAGTGA	693
Db	421	GAAGAAACCGATTTGGACACAGTCGGAAATGCTTCTTACGATATACCAAGAGAGTGA	480
Qy	694	ACATTTCTGTTTCAAGCCGGTAGTGGAAATTTATCAAGTAAAGATGGAGGCCACAGGA	753
Db	481	ACATTTCTGTTTCAAGCCGGTAGTGGAAATTTATCAAGTAAAGATGGAGGCCACAGGA	540
Qy	754	TTTACGCAACAACTTTAAGGCCCAATCTAGTGGAACTAGTTGTCACACATACCGATG	813
Db	541	TTTACGCAACAACTTTAAGGCCCAATCTAGTGGAACTAGTTGTCACACATACCGATG	600
Qy	814	GATCCAAAATTTATGCCCTGCTGATCCAGACTGGATTTGCTTTTATACATAGCAAGATTT	873
Db	601	GATCCAAAATTTATGCCCTGCTGATCCAGACTGGATTTGCTTTTATACATAGCAAGATTT	560
Qy	874	TGGATATCTAACTCGTAAACAGAGAGAAAGGAGACTCACTTATATGCAATATGAGCTA	933
Db	661	TGGATATCTAACTCGTAAACAGAGAGAAAGGAGACTCACTTATATGCAATATGAGCTA	720
Qy	934	GCCAATGGAAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAGAGAA	993
Db	721	GCCAATGGAAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAGAGAA	780
Qy	994	TTTGATAGATTTCTGGCTATTGGTGTGCCAAAGCTGAAACAACTCCAGTGGTGGT	1053
Db	781	TTTGATAGATTTCTGGCTATTGGTGTGCCAAAGCTGAAACAACTCCAGTGGTGGT	840
Qy	1054	AAAATCTTGAATCTCTATATGAAGAAATGATGAATCTGAGGTGGAAATTTATTCATGTT	1113
Db	841	AAAATCTTGAATCTCTATATGAAGAAATGATGAATCTGAGGTGGAAATTTATTCATGTT	900
Qy	1114	ACATCCCTATGTTGGAAACAGAGGGGCAGATTCAATCCGTTATCCCTAAACACAGGTACA	1173

901	ACATCCCTATGTTGGAAACAGAGGGGCAGATTCAATTCGTTATCCCTAAACACAGGTACA	960
1174	GCAATCTTAAAGTCACTTTTAAAGTGTGAGAAATTAATGATTTGATGCTGAAAGGAAGATC	1233
961	GCAATCTTAAAGTCACTTTTAAAGTGTGAGAAATTAATGATTTGATGCTGAAAGGAAGATC	1020
1234	ATAGATGTCTAGATTAAGGAACCTAATTTCAACCTTTTGGATTTCTATTTGAAAGGAGTTGAA	1293
1021	ATAGATGTCTAGATTAAGGAACCTAATTTCAACCTTTTGGATTTCTATTTGAAAGGAGTTGAA	1080
1294	TATATTTGCCAGAGCTGGATGGAATCTCTGAGGGGAAATATGCTTGGCTCCTACTACATAGT	1353
1081	TATATTTGCCAGAGCTGGATGGAATCTCTGAGGGGAAATATGCTTGGCTCCTACTACATAGT	1140
1354	CGCTCCCAAGCTCGCCCTACAGATAGTGTGATCTCACTGAAATTTATTTATCCAGTAGAA	1413
1141	CGCTCCCAAGCTCGCCCTACAGATAGTGTGATCTCACTGAAATTTATTTATCCAGTAGAA	1200
1414	GATGATGTTATGGAAGGACAGAGACTCATTTGAGTCAAGTCCCTGATTTCTGTGACGCACATA	1473
1201	GATGATGTTATGGAAGGACAGAGACTCATTTGAGTCAAGTCCCTGATTTCTGTGACGCACATA	1260
1474	ATTATCTATGAAAGAAACACAGACATCTGGATTAATATCCATGACATCTTTTCAATGTTTTT	1533
1261	ATTATCTATGAAAGAAACACAGACATCTGGATTAATATCCATGACATCTTTTCAATGTTTTT	1320
1534	CCCCAAAGCTCAGGAAGAGGAAATTTGAGTTTATTTTGGCTCTGAAATGCAAAAACAGGTTTC	1593
1321	CCCCAAAGCTCAGGAAGAGGAAATTTGAGTTTATTTTGGCTCTGAAATGCAAAAACAGGTTTC	1380
1594	CGTCAATTTATCAAAATTTACATCTATTTTAAAGGAAGCAAAATATAAAGCATGCTGAGTGT	1653
1381	CGTCAATTTATCAAAATTTACATCTATTTTAAAGGAAGCAAAATATAAAGCATGCTGAGTGT	1440
1654	GGGCTGCTGCTCCAGAGTATTTCAAGTGTCTTATCAAGAGGAGATAGCAATTTACAGT	1713
1441	GGGCTGCTGCTCCAGAGTATTTCAAGTGTCTTATCAAGAGGAGATAGCAATTTACAGT	1500
1714	GGTGAATGGAAGTTCTTTGGCCGGCATGGAATTAATATCCAACTGATGAAGTCAGAAG	1773
1501	GGTGAATGGAAGTTCTTTGGCCGGCATGGAATTAATATCCAACTGATGAAGTCAGAAG	1560
1774	CTGGTATATTTTGAAGGACCAAAAGACTCCCTTTTAGAGCATCCCTGTACGTAGTCAGT	1833
1561	CTGGTATATTTTGAAGGACCAAAAGACTCCCTTTTAGAGCATCCCTGTACGTAGTCAGT	1620
1834	TACGTAATCTCTGGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCAGATTTCTGCTGC	1893
1621	TACGTAATCTCTGGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCAGATTTCTGCTGC	1680
1894	ATCAGTCAGCAGCTGTGACTTTCTTTATTAAGTAAGTATAGTAAACAGAGAAATCCACTGT	1953
1681	ATCAGTCAGCAGCTGTGACTTTCTTTATTAAGTAAGTATAGTAAACAGAGAAATCCACTGT	1740
1954	GTGTCCCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAGAAAACAAAGGAAATTT	2013
1741	GTGTCCCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCAGAAAACAAAGGAAATTT	1800
2014	TGGGCCACCAATTTTGGATTTGAGTTCAGCAGGTCTCTTCTGCTGACTACTCTCCCTCAGAAAATTTTC	2073
1801	TGGGCCACCAATTTTGGATTTGAGTTCAGCAGGTCTCTTCTGCTGACTACTCTCCCTCAGAAAATTTTC	1860
2074	TCCTTTGAAAGTACTACTGATTTTACATTTGATGGATGCTCTCAAGAGCTCTCATGATCTA	2133
1861	TCCTTTGAAAGTACTACTGATTTTACATTTGATGGATGCTCTCAAGAGCTCTCATGATCTA	1920
2134	CAGCTGGAAAGAAATATCTACTGCTGCTTATATATATGTTGGTCTCTCAGGTGAGTTG	2193
1921	CAGCTGGAAAGAAATATCTACTGCTGCTTATATATATGTTGGTCTCTCAGGTGAGTTG	1980
2194	GTGAATTAATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCCTCAGCTCTCTAGGT	2253

Db 1981 GTGAATAATCGGTTAAAGAGTCAAGTATTTCCGGTTGAATAACCTAGCCTCTCTAGGT 2040
QY 2254 TATGTGTTGTAGTAGACAAACAGGGGATCCTGTCCACCGAGGCTTAAATTTGAAGGC 2313
Db 2041 TATGTGTTGTAGTAGACAAACAGGGGATCCTGTCCACCGAGGCTTAAATTTGAAGGC 2100
QY 2314 GCCTTTAAATATAAATGGGTCAAAATAGAAATTAAGATCAGGTGAAGACTCCAATAT 2373
Db 2101 GCCTTTAAATATAAATGGGTCAAAATAGAAATTAAGATCAGGTGAAGACTCCAATAT 2160
QY 2374 CTAGCTCTCCGATATGATTTTCATTGACTTAGATCGTGTGGCATCCACGGCTGCTCTAT 2433
Db 2161 CTAGCTCTCCGATATGATTTTCATTGACTTAGATCGTGTGGCATCCACGGCTGCTCTAT 2220
QY 2434 GGAGGATACCTCTCCCTGATGGCATTAATGCAGAGTCAAGATCTTTCAGGGTTGCTATT 2493
Db 2221 GGAGGATACCTCTCCCTGATGGCATTAATGCAGAGTCAAGATCTTTCAGGGTTGCTATT 2280
QY 2494 GCTGGGGCCCGAGTCACCTCTGTGGATCTTCTATGATACAGGATACACGGAACTTATATG 2553
Db 2281 GCTGGGGCCCGAGTCACCTCTGTGGATCTTCTATGATACAGGATACACGGAACTTATATG 2340
QY 2554 GGTACACCTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCATGCAAGCAGAA 2613
Db 2341 GGTACACCTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCATGCAAGCAGAA 2400
QY 2614 AAGTTCCCTCTGAACCAAAATCGTTTACTGTCTTTACATGGTTTCTCTGGATGAGATGTC 2673
Db 2401 AAGTTCCCTCTGAACCAAAATCGTTTACTGTCTTTACATGGTTTCTCTGGATGAGATGTC 2460
QY 2674 CATTTTGCACATACCAAGTATATTACTGAGTTTTTTTAGTGGGCTCGAAAGCCATATGAT 2733
Db 2461 CATTTTGCACATACCAAGTATATTACTGAGTTTTTTTAGTGGGCTCGAAAGCCATATGAT 2520
QY 2734 TTACAGATCTATCTCTCAGGAGAGACAGACATAGAGTTCTCGAATCGGGAGACATTTAT 2793
Db 2521 TTACAGATCTATCTCTCAGGAGAGACAGACATAGAGTTCTCGAATCGGGAGACATTTAT 2580
QY 2794 GAACCTGCATCTTTTGCATCTACCTTCAAGAAAACCTTGGATCAGTATTTGCTCTTAAAA 2853
Db 2581 GAACCTGCATCTTTTGCATCTACCTTCAAGAAAACCTTGGATCAGTATTTGCTCTTAAAA 2640
QY 2854 GTGATATAA 2862
Db 2641 GTGATATAA 2649

Search completed: April 15, 2006, 00:03:00
Job time : 15361 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 14, 2006, 19:20:55 ; Search time 1768 Seconds
(without alignments)
11761.222 Million cell updates/sec

Title: US-10-825-632-2
Perfect score: 3120
Sequence: 1 aagtgctaaagctctcgag.....agaattactaaaaaaaaa 3120

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_21.*

- 1: geneseq1980s.*
- 2: geneseq1990s.*
- 3: geneseq2000s.*
- 4: geneseq2001as.*
- 5: geneseq2001bs.*
- 6: geneseq2002as.*
- 7: geneseq2002bs.*
- 8: geneseq2003as.*
- 9: geneseq2003bs.*
- 10: geneseq2003cs.*
- 11: geneseq2003ds.*
- 12: geneseq2004as.*
- 13: geneseq2004bs.*
- 14: geneseq2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3120	100.0	3120	4 AAC85694	Aac85694 Nucleotid
2	3120	100.0	3120	6 AAD38956	Aad38956 Human dip
3	3106.4	99.6	4829	6 ABK83327	Abk83327 cDNA enco
4	3105.2	99.5	3143	6 AAH99934	Aah99934 cDNA enco
5	3101.8	99.4	3106	6 ABK12892	Abk12892 Human pro
6	2824.4	90.5	4685	6 ABK83332	Abk83332 cDNA enco
7	2806.4	89.9	4676	6 ABK83331	Abk83331 cDNA enco
8	2785.8	89.3	2952	10 ACA92421	Ac92421 DNA enco
9	2759.4	88.4	2797	12 ADL13374	Adl13374 Human ste
10	2742.4	87.9	2929	10 ACA92425	Ac92425 DNA enco
11	2647.4	84.9	2649	14 ADV43981	Adv43981 Human psy
12	2644.4	84.8	2671	6 ABK83322	Abk83322 cDNA enco
13	2644.2	84.7	2649	8 ABK12255	Abk12255 cDNA enco
14	2639.8	84.6	2643	6 AAH99935	Aah99935 Coding se
15	2500.4	80.1	4523	6 ABK83325	Abk83325 cDNA enco
16	2490	79.8	2842	6 ABN59774	Abn59774 Novel hum
17	2315.8	74.2	2696	12 ADI16386	Adi16386 Human pro
18	2208	70.8	2510	6 AAD23843	Aad23843 Human pro
19	2167.4	69.5	2702	13 ADT04072	Adt04072 Human pro

20	2132	68.3	2668	6 ABN59775	Abn59775 Novel hum
21	2056.4	65.9	4309	6 ABK83328	Abk83328 cDNA enco
22	2037.4	65.3	2349	14 ADV43982	Adv43982 Human psy
23	1994.8	63.9	2161	4 AAH15009	Aah15009 Human cDN
c	24	1994	2083	7 ADR41222	Adr41222 Human CD-
25	1827	58.6	2463	10 ACA92424	Ac92424 DNA enco
26	1819.4	58.3	1821	6 ABV76411	Abv76411 Dipeptid
27	1774.8	56.9	2251	10 ADE79035	Ad79035 Human pro
28	1347	43.2	1669	4 AAC85696	Aac85696 Nucleotid
29	1036	33.2	1197	4 AAC85695	Aac85695 Nucleotid
30	926.4	29.7	1356	6 ABK83326	Abk83326 cDNA enco
31	883.8	28.3	3287	6 AAD38955	Aad38955 Alternati
32	862.6	27.6	2660	6 ADI16689	Adi16689 Human NOV
33	862.6	27.6	2660	6 ADI16687	Adi16687 Human NOV
34	861	27.6	2617	6 ABK83323	Abk83323 cDNA enco
35	861	27.6	2660	12 ADN42341	Adn42341 Human cDN
36	861	27.6	3716	6 ABQ75955	Abq75955 Human PMM
37	861	27.6	4147	12 ADQ86753	Adq86753 Human tum
38	861	27.6	4147	13 ADQ84273	Adq84273 Human tum
39	861	27.6	4147	13 ACN40423	Acn40423 Tumour-as
40	861	27.6	4219	6 ABK83335	Abk83335 cDNA enco
41	861	27.6	4302	6 ABK83333	Abk83333 cDNA enco
42	861	27.6	4398	13 ADS10267	Ads10267 Human the
43	857.8	27.5	3024	6 AAD38954	Aad38954 Human dip
44	854	27.4	2495	6 AAD38957	Aad38957 Human dip
45	840	26.9	2751	6 AAD38311	Aad38311 Murine di

ALIGNMENTS

RESULT 1
AAC85694

ID AAC85694 standard; cDNA; 3120 BP.

XX AC AAC85694;

XX DT 29-JUN-2001 (first entry)

XX DE Nucleotide sequence of human DPP8.

XX HW Human; dipeptidyl aminopeptidase; DPP8; prolyl oligopeptidase;
XX KW dipeptidyl peptidase; DPPIV; T cell; cleavage; diarrhoea;
XX KW growth hormone deficiency; glucose level; mucosal regeneration;
XX KW non-insulin dependent diabetes mellitus; glucose intolerance;
XX KW immunosuppression; ss.

XX OS Homo sapiens.

XX Key Location/Qualifiers
XX CDS 214..2862

XX FT /*tag= a

XX FT /product= "Human DPP8"

XX FN WO200119866-A1.

XX PD 22-MAR-2001.

XX PF 11-SEP-2000; 2000WO-AU001085.

XX PR 10-SEP-1999; 99AU-00002762.

XX PR 18-FEB-2000; 2000AU-00005709.

XX PA (UNSY) UNIV SYDNEY.

XX PI Abbott CA, Gorell MD;

XX DR WPI; 2001-281520/29.

XX DR P-PSDB; AAB47187.

XX FT New human dipeptidyl aminopeptidase (DPP8) useful for cleaving

XX FT substrates, identifying inhibitors of DPP8 catalytic activity which have

XX FT therapeutic uses, and for detecting activated T cells.

XX Claim 16; Fig 2; 78pp; English.

XX This sequence encodes human dipeptidyl aminopeptidase (DPP8). DPP8 has

CC substrate specificity for H-Ala-Pro-pNA, H-Gly-Pro-pNA and H-Arg-Pro-pNA.

CC Therefore, it is a prolyl oligopeptidase and a dipeptidyl peptidase,

CC because it is capable of hydrolysing the peptide bond C-terminal to Pro

CC in each of these compounds. DPP8 is homologous with human DPPIV. DPP8 is

CC useful for cleaving a substrate, and for detecting an activated T cell

CC which involves measuring the level of DPP8 gene expression in a T cell.

CC The level of DPP8 expression is detected by detecting the amount of DPP8

CC RNA in the cell. It is also useful for identifying a molecule capable of

CC inhibiting the cleavage of the substrate by DPP8. Molecules identified as

CC inhibiting DPP8 catalytic activity may be useful for treating diarrhoea,

CC growth hormone deficiency, lowering glucose levels in non-insulin

CC dependent diabetes mellitus and other disorders involving glucose

CC intolerance, enhancing mucosal regeneration and as immunosuppressants

XX Sequence 3120 BP; 936 A; 637 C; 706 G; 841 T; 0 U; 0 Other;

SQ

Query Match 100.0%; Score 3120; DB 4; Length 3120;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 3120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGTCTTAAGCCTCCGAGGCCAAGCGCGCTGTACTGCGCGCGCTTCTTACTGCGG 60

DB 1 AAGTCTTAAGCCTCCGAGGCCAAGCGCGCTGTACTGCGCGCGCTTCTTACTGCGG 60

QY 61 CGTTCCGCGCCTTGGGTTGTCAACGCGCGCGCGCGAGGAGGCCATGCAACACGAGACG 120

DB 61 CGTTCCGCGCCTTGGGTTGTCAACGCGCGCGCGCGAGGAGGCCATGCAACACGAGACG 120

QY 121 GAGTGGAGCGCGCAGCATGAAGCGCGCAGCGCGCTCCATAGCGCACGTCGGGACGG 180

DB 121 GAGTGGAGCGCGCAGCATGAAGCGCGCAGCGCGCTCCATAGCGCACGTCGGGACGG 180

QY 181 TCCGCGCGCGCGCGCGGGAAGGAAATGCAATATGCGCAGCAGCAATGGAACAGAACAG 240

DB 181 TCCGCGCGCGCGCGGGAAGGAAATGCAATATGCGCAGCAGCAATGGAACAGAACAG 240

QY 241 CTGGGTGTGGAGATATTGAAATCGCGAGCTGTGAGGAGAAATTCGAATCAGAGATCGG 300

DB 241 CTGGGTGTGGAGATATTGAAATCGCGAGCTGTGAGGAGAAATTCGAATCAGAGATCGG 300

QY 301 CCTAAATGGAGCCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGAGCTGCTT 360

DB 301 CCTAAATGGAGCCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGAGCTGCTT 360

QY 361 GCGGATACGAGAAATATCATGCTACATGATGCTTAAGGACCAACATCATATTCATGTTT 420

DB 361 GCGGATACGAGAAATATCATGCTACATGATGCTTAAGGACCAACATCATATTCATGTTT 420

QY 421 GTGAAGAGATGATCCAGATGGAGCTCATTCAGACAGATCTATTACCTGCCATGTCT 480

DB 421 GTGAAGAGATGATCCAGATGGAGCTCATTCAGACAGATCTATTACCTGCCATGTCT 480

QY 481 GGTGAGACAGAGAAAATACACTGTGTTTATCTGAAATTCGAAATCTATCATATAGACA 540

DB 481 GGTGAGACAGAGAAAATACACTGTGTTTATCTGAAATTCGAAATCTATCATATAGACA 540

QY 541 GCAGTCTTAATGCTCTCTTGAAGCCTCTTTTGGATCTTTTTCAGGCAACATCTGGACTAT 600

DB 541 GCAGTCTTAATGCTCTCTTGAAGCCTCTTTTGGATCTTTTTCAGGCAACATCTGGACTAT 600

QY 601 GGAATGTATTCGAGAGAGAGAACTATTAAAGAGAAAGAAACGCAATGGAACAGTCGGA 660

DB 601 GGAATGTATTCGAGAGAGAGAACTATTAAAGAGAAAGAAACGCAATGGAACAGTCGGA 660

QY 661 ATTGCTTCTTACGATTATCACCAGGAAGTGAACATTTCTGTTTCAAGCCGGTAGTGA 720

DB 661 ATTGCTTCTTACGATTATCACCAGGAAGTGAACATTTCTGTTTCAAGCCGGTAGTGA 720

QY 721 ATTTATACGTAAAGATGGAGGGCCACAGGATTTTACGCAACAACTTTTAAAGGCCCAAT 780

Db QY 721 ATTTATACGTAAAGATGGAGGGCCACAGGATTTTACGCAACAACTTTTAAAGGCCCAAT 780

QY 781 CTAGTGGAAAACCTAGTTGTCCCAACATACGATGATGATCCCAAAATATATCCCGCTGATCA 840

DB 781 CTAGTGGAAAACCTAGTTGTCCCAACATACGATGATGATCCCAAAATATATCCCGCTGATCA 840

QY 841 GACTCGATGCTGCTTTTATACATAGCAACGATATTGAGATATCTAACATCGTAAACAGAGAA 900

DB 841 GACTCGATGCTGCTTTTATACATAGCAACGATATTGAGATATCTAACATCGTAAACAGAGAA 900

QY 901 GAAAGGAGACTCCTTATGTGTCACATAGCTAGCCAAACATGGAAGAAAGATGCCAGATCA 960

DB 901 GAAAGGAGACTCCTTATGTGTCACATAGCTAGCCAAACATGGAAGAAAGATGCCAGATCA 960

QY 961 GCTGAGTGGCTTACCTTTGTTCTCCAAAGAAAGATTTGATAGATATTTCTGCTATTGGTGG 1020

DB 961 GCTGAGTGGCTTACCTTTGTTCTCCAAAGAAAGATTTGATAGATATTTCTGCTATTGGTGG 1020

QY 1021 TGTCCAAAAGCTGAAAACAACTCCCAAGTGGTGGTAAAAATTTCTTAGAAATTTCTATATGAAGAA 1080

DB 1021 TGTCCAAAAGCTGAAAACAACTCCCAAGTGGTGGTAAAAATTTCTTAGAAATTTCTATATGAAGAA 1080

QY 1081 AATGATGAATCTGAGGTGGAAATTTATTCATGTTTACATCCCTCTATGTTGGAACAAAGGAGG 1140

DB 1081 AATGATGAATCTGAGGTGGAAATTTATTCATGTTTACATCCCTCTATGTTGGAACAAAGGAGG 1140

QY 1141 GCAGATTCATTTCCGTTATCTTAAACAGGTACAGCAATCTTAAAGTCACCTTTTAAAGATG 1200

DB 1141 GCAGATTCATTTCCGTTATCTTAAACAGGTACAGCAATCTTAAAGTCACCTTTTAAAGATG 1200

QY 1201 TCGAAATATGATTTGATGCTGAGGAGGATCATAGATGCTCATAGTATAGGAACCTAATTT 1260

DB 1201 TCGAAATATGATTTGATGCTGAGGAGGATCATAGATGCTCATAGTATAGGAACCTAATTT 1260

QY 1261 CAACTTTTGAAGTCTATTTTGAAGGAGTTGAAATATATTCGCGAGAGCTGATGAGACTCT 1320

DB 1261 CAACTTTTGAAGTCTATTTTGAAGGAGTTGAAATATATTCGCGAGAGCTGATGAGACTCT 1320

QY 1321 GAGGGAATATGCTTGGTCCATCTACTAGATCGCTCCAGACTCGCTTACAGATAGTG 1380

DB 1321 GAGGGAATATGCTTGGTCCATCTACTAGATCGCTCCAGACTCGCTTACAGATAGTG 1380

QY 1381 TTGATCTCACCTGAAATATTTATCCAGTAGAAGATGATGTTATGGAAGGCGAGAGACTC 1440

DB 1381 TTGATCTCACCTGAAATATTTATCCAGTAGAAGATGATGTTATGGAAGGCGAGAGACTC 1440

QY 1441 ATTGAGTCACTGCTGATTTCTGTGACGCCACTAATTTATCTATGAAGAAACAAACAGACATC 1500

DB 1441 ATTGAGTCACTGCTGATTTCTGTGACGCCACTAATTTATCTATGAAGAAACAAACAGACATC 1500

QY 1501 TGGATAAATATCCATGACATCTTTTCATGTTTCCCAAGTCCAGAGAGGAAATTTGAG 1560

DB 1501 TGGATAAATATCCATGACATCTTTTCATGTTTCCCAAGTCCAGAGAGGAAATTTGAG 1560

QY 1561 TTTATTTTGGCTCTGAAATGCAAAACAGGTTTCCGTCAATTTATACAAAATTTACATCTATT 1620

DB 1561 TTTATTTTGGCTCTGAAATGCAAAACAGGTTTCCGTCAATTTATACAAAATTTACATCTATT 1620

QY 1621 TTAAGGAAACAAATATATAAGCATCCAGTGGCTGCTGCTCCAGTGGATTTTCAAG 1680

DB 1621 TTAAGGAAACAAATATATAAGCATCCAGTGGCTGCTGCTCCAGTGGATTTTCAAG 1680

QY 1681 TGTCTCTATCAAGAGAGATAGCAATTTACAGTGGTGAATGGAGGTTCTTGGCCGCAT 1740

DB 1681 TGTCTCTATCAAGAGAGATAGCAATTTACAGTGGTGAATGGAGGTTCTTGGCCGCAT 1740

QY 1741 GGATCTAATATCCAAAGTTGATGAGTCAAGAGGCTGGTATATTTTGAAGGCCAACAGAC 1800

DB 1741 GGATCTAATATCCAAAGTTGATGAGTCAAGAGGCTGGTATATTTTGAAGGCCAACAGAC 1800

QY 1801 TCCCTCTTAGAGCATCCTGTACCTAGTCTAGTTACGTAATCGTAAATCTCTGGAGGTTGCAAGG 1860

QY 1 AAGTGCTAAGCCCTCCGAGGCCAAGCCCGCTGCTACTGCCGCGCTCTTCTTAGTGCGG 60

Db 1 AAGTCCTAAAGCCCTCCGAGGCCAAGGCCGCTGCTACTGCGCGCGCTGCTTCTTAGTGCGG 60
Qy 61 GCTTCGCGCGCTGGGTTGTCAACGCGCGCGCGCGCGAGGAAGCACTGCAACACGAGACGG 120
Db 61 GCTTCGCGCGCTGGGTTGTCAACGCGCGCGCGCGCGAGGAAGCACTGCAACACGAGACGG 120
Qy 121 GAGTGGAGCGCGCAGCATGAAGCGCGCAGGCGCGCTCCATAGCGCACTGTCGGGACGG 180
Db 121 GAGTGGAGCGCGCAGCATGAAGCGCGCAGGCGCGCTCCATAGCGCACTGTCGGGACGG 180
Qy 181 TCCGCGCGCGCGCGCGGGAAGGAATAATGCAACATGCGCAGCAGCAATGGAACAGAACAG 240
Db 181 TCCGCGCGCGCGCGCGGGAAGGAATAATGCAACATGCGCAGCAGCAATGGAACAGAACAG 240
Qy 241 CTGGGTGTGAGATATTGAAACTGCGGACTGTGAGGAGATAATTGAATCAGAGATCGG 300
Db 241 CTGGGTGTGAGATATTGAAACTGCGGACTGTGAGGAGATAATTGAATCAGAGATCGG 300
Qy 301 CCTAAATTGGAGCCCTTTTATGTTGAGCGGTATTCTGGAGTCAGCTTAAAGCTGCTT 360
Db 301 CCTAAATTGGAGCCCTTTTATGTTGAGCGGTATTCTGGAGTCAGCTTAAAGCTGCTT 360
Qy 361 GCCGATACAGAAATATATGCTCATGATGCTTAGGCCACCAATGATTTCAATGTTT 420
Db 361 GCCGATACAGAAATATATGCTCATGATGCTTAGGCCACCAATGATTTCAATGTTT 420
Qy 421 GTGAAGAGGAATGATCCAGATGGAACCTCATTCAGACAGAAATCTATTACCTTGCCATGTCT 480
Db 421 GTGAAGAGGAATGATCCAGATGGAACCTCATTCAGACAGAAATCTATTACCTTGCCATGTCT 480
Qy 481 GGTGAGAAACAGAGAAATACACTGTGTTTATCTGAAATTTCCCAAAACTATATCAATAGACA 540
Db 481 GGTGAGAAACAGAGAAATACACTGTGTTTATCTGAAATTTCCCAAAACTATATCAATAGACA 540
Qy 541 GCAGTCTTAATGCTCTCTTGGAAGCCTCTTTTGGATCTTTTTCAGGCAACACCTGGACTAT 600
Db 541 GCAGTCTTAATGCTCTCTTGGAAGCCTCTTTTGGATCTTTTTCAGGCAACACCTGGACTAT 600
Qy 601 GGAATGTATCTCGAGAAGAGAACTATTAAAGAGAAAGAAACGCAATGGAAACAGTCGGA 660
Db 601 GGAATGTATCTCGAGAAGAGAACTATTAAAGAGAAAGAAACGCAATGGAAACAGTCGGA 660
Qy 661 ATTTGCTCTTACGATTATACCAAGGAAGTGAACATTTCTGTTCAAGCCGCTAGTGA 720
Db 661 ATTTGCTCTTACGATTATACCAAGGAAGTGAACATTTCTGTTCAAGCCGCTAGTGA 720
Qy 721 ATTTATCAGTAAGATGGAGGCCACAGGATTTACGCAACGACCTTTAAGGCCCAAT 780
Db 721 ATTTATCAGTAAGATGGAGGCCACAGGATTTACGCAACGACCTTTAAGGCCCAAT 780
Qy 781 CTAGTGGAAAACCTAGTGTGCCAATACGGAATGATCGAAATTTATGCCCCCGCTGATCCA 840
Db 781 CTAGTGGAAAACCTAGTGTGCCAATACGGAATGATCGAAATTTATGCCCCCGCTGATCCA 840
Qy 841 GACTGGATTGCTTTTATACATAGCAACGATATTGGATATCTAAATCGTAAACGAGAA 900
Db 841 GACTGGATTGCTTTTATACATAGCAACGATATTGGATATCTAAATCGTAAACGAGAA 900
Qy 901 GAAAGAGACTCACTTATGTGCAACATGAGTGAAGCAACATGGAAGAGATGCCAGATCA 960
Db 901 GAAAGAGACTCACTTATGTGCAACATGAGTGAAGCAACATGGAAGAGATGCCAGATCA 960
Qy 961 GCTGAGTGCCTACCTTTGTTCTCAAGAGAAATTTGATAGATATTCTGGCTATTGGTGG 1020
Db 961 GCTGAGTGCCTACCTTTGTTCTCAAGAGAAATTTGATAGATATTCTGGCTATTGGTGG 1020
Qy 1021 TGTCCAAAAGCTGAAACAACTCCAGTGGTGGTAAATTTCTTAGAATTTCTATATGAAGA 1080
Db 1021 TGTCCAAAAGCTGAAACAACTCCAGTGGTGGTAAATTTCTTAGAATTTCTATATGAAGA 1080
Qy 1081 AATGATGATCTGAGGTGGAATTTATTCATGTTTACATCCCTATGTTGGAACAGAGG 1140
Db 1081 AATGATGATCTGAGGTGGAATTTATTCATGTTTACATCCCTATGTTGGAACAGAGG 1140

Qy 1141 GCAGATTCTTCGTTATCTCTAAACAGGTTACAGCAATCCTAAAGTCACTTTTAAAGTG 1200
Db 1141 GCAGATTCTTCGTTATCTCTAAACAGGTTACAGCAATCCTAAAGTCACTTTTAAAGTG 1200
Qy 1201 TCAGAAAATAATGATTGATGCTGAAAGGAAGGATCATAGATGCTATAGATAAGGAACATAAT 1260
Db 1201 TCAGAAAATAATGATTGATGCTGAAAGGAAGGATCATAGATGCTATAGATAAGGAACATAAT 1260
Qy 1261 CAACCTTTTGAGATCTATTGAAAGGAGTTGAATATATTGCCAGAGCTGATCGACTCCT 1320
Db 1261 CAACCTTTTGAGATCTATTGAAAGGAGTTGAATATATTGCCAGAGCTGATCGACTCCT 1320
Qy 1321 GAGGGAATAATGCTTGGTCCATCTACTAGATCGCTCCAGAGCTCGCTTACAGATAGTG 1380
Db 1321 GAGGGAATAATGCTTGGTCCATCTACTAGATCGCTCCAGAGCTCGCTTACAGATAGTG 1380
Qy 1381 TTGATCTCACCTGAAATATTATCCAGTAGAAGATGATGTTATGGAAGGCGAGAGACTC 1440
Db 1381 TTGATCTCACCTGAAATATTATCCAGTAGAAGATGATGTTATGGAAGGCGAGAGACTC 1440
Qy 1441 ATTTGAGTCAGTCCCTGATTTCTGACGCCACTAATTTATCTATGAAGAAAACACAGACATC 1500
Db 1441 ATTTGAGTCAGTCCCTGATTTCTGACGCCACTAATTTATCTATGAAGAAAACACAGACATC 1500
Qy 1501 TGGATAAATATCCATGACACTCTTTTCATGTTTCCCAAAGTCAAGAGGAAATTTGAG 1560
Db 1501 TGGATAAATATCCATGACACTCTTTTCATGTTTCCCAAAGTCAAGAGGAAATTTGAG 1560
Qy 1561 TTTATTTTGGCTCTGAAATGCAAAACAGGTTTCCGTCATTTATACAAAATTAATCATCTATT 1620
Db 1561 TTTATTTTGGCTCTGAAATGCAAAACAGGTTTCCGTCATTTATACAAAATTAATCATCTATT 1620
Qy 1621 TTAAGGAAAGCAAAATATAAAGATCCAGTGGTGGCTGCTGCTCCAAAGTGAATTTCAAG 1680
Db 1621 TTAAGGAAAGCAAAATATAAAGATCCAGTGGTGGCTGCTGCTCCAAAGTGAATTTCAAG 1680
Qy 1681 TGTCTCTATCAAAAGAGGAGATAGCAATTTACAGTGGTGAATGGGAAGTTCTTGGCCGCGAT 1740
Db 1681 TGTCTCTATCAAAAGAGGAGATAGCAATTTACAGTGGTGAATGGGAAGTTCTTGGCCGCGAT 1740
Qy 1741 GGATCTAATATCCAAAGTTGATGAAGTCAAGAGGCTGGTATATTTTGAAGGCAACCAAGAC 1800
Db 1741 GGATCTAATATCCAAAGTTGATGAAGTCAAGAGGCTGGTATATTTTGAAGGCAACCAAGAC 1800
Qy 1801 TCCCTTTTAGAGCATCACCTGTAGTAGTACGTTACGTTAAATCCTGGAGAGGTGACAAGG 1860
Db 1801 TCCCTTTTAGAGCATCACCTGTAGTAGTACGTTAAATCCTGGAGAGGTGACAAGG 1860
Qy 1861 CTGACTGACCGTGGCTACTCACTTCTGCTGATCAGTCAGCACTGCTGACTTCTTTATA 1920
Db 1861 CTGACTGACCGTGGCTACTCACTTCTGCTGATCAGTCAGCACTGCTGACTTCTTTATA 1920
Qy 1921 AGTAAGTATAGTAACAGAGAAATCCACACTGTGTGCTCCCTTTTACAGCTATCAAGTCT 1980
Db 1921 AGTAAGTATAGTAACAGAGAAATCCACACTGTGTGCTCCCTTTTACAGCTATCAAGTCT 1980
Qy 1981 GAAGATGACCCAACTTGCAAAACAAAGGAATTTTGGGCCACATTTTGGATTTGAGAGGT 2040
Db 1981 GAAGATGACCCAACTTGCAAAACAAAGGAATTTTGGGCCACATTTTGGATTTGAGAGGT 2040
Qy 2041 CCTCTTCTGACTATACCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGAATTTACA 2100
Db 2041 CCTCTTCTGACTATACCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGAATTTACA 2100
Qy 2101 TTGATAGGATGCTCTACAAGCTCATGATCTACAGCTCGGAAAGAAATATCTCTACTGTG 2160
Db 2101 TTGATAGGATGCTCTACAAGCTCATGATCTACAGCTCGGAAAGAAATATCTCTACTGTG 2160
Qy 2161 CTGTTTCATATATGTTGGTCTCAGGTGAGTTGGTGAATATCGTTTTTAAAGAGGTCAAG 2220
Db 2161 CTGTTTCATATATGTTGGTCTCAGGTGAGTTGGTGAATATCGTTTTTAAAGAGGTCAAG 2220

QY 2221 TATTTCCGCTTGAATACCTAGCTCTCTAGGTTATGTGGTTGTAGTAGAACAACAGG 2280
Db 2221 TATTTCCGCTTGAATACCTAGCTCTCTAGGTTATGTGGTTGTAGTAGAACAACAGG 2280
QY 2281 GATCTCTGTACCCAGGGCTTAAATTTGAAGGCGCTTTAAATATATAAATATGAGTCAATA 2340
Db 2281 GATCTCTGTACCCAGGGCTTAAATTTGAAGGCGCTTTAAATATATAAATATGAGTCAATA 2340
QY 2341 GAAATTTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTCAATTGAC 2400
Db 2341 GAAATTTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTCAATTGAC 2400
QY 2401 TTAGATCGTGTGGGCATCCAGCGTGTCTTATGAGGATACCTCTCCCTGATGCGATTA 2460
Db 2401 TTAGATCGTGTGGGCATCCAGCGTGTCTTATGAGGATACCTCTCCCTGATGCGATTA 2460
QY 2461 ATGCGAGGTGAGATATCTTCAGGGTTGCTATTTGCTGGGGCCCGCAGTCACTCTGTGGATC 2520
Db 2461 ATGCGAGGTGAGATATCTTCAGGGTTGCTATTTGCTGGGGCCCGCAGTCACTCTGTGGATC 2520
QY 2521 TTCTATGATACAGGATACAGGAACGTTATATGGGTCACTGACAGGAATGAACAGGGC 2580
Db 2521 TTCTATGATACAGGATACAGGAACGTTATATGGGTCACTGACAGGAATGAACAGGGC 2580
QY 2581 TATTACTTATGAGTCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGAAACCAATCGTTTA 2640
Db 2581 TATTACTTATGAGTCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGAAACCAATCGTTTA 2640
QY 2641 CTGCTCTTACATGGTTCTGGATGAGATGTCTCATTTTGCATACCAATGATATCTG 2700
Db 2641 CTGCTCTTACATGGTTCTGGATGAGATGTCTCATTTTGCATACCAATGATATCTG 2700
QY 2701 AGTTTTTTAGTGGGCTGGAAGCCATATGATTTACAGATCTATCTCAGAGAGACAC 2760
Db 2701 AGTTTTTTAGTGGGCTGGAAGCCATATGATTTACAGATCTATCTCAGAGAGACAC 2760
QY 2761 AGCATAAGAGTTCTTGAATCGGAGAACATTTATGAATGCAATCTTTTGCATACCTTCAA 2820
Db 2761 AGCATAAGAGTTCTTGAATCGGAGAACATTTATGAATGCAATCTTTTGCATACCTTCAA 2820
QY 2821 GAAACCTTTGGATCACTGATTTGCTGCTCTAAAGTATATAATTTGACCTGTGTAGAAC 2880
Db 2821 GAAACCTTTGGATCACTGATTTGCTGCTCTAAAGTATATAATTTGACCTGTGTAGAAC 2880
QY 2881 TCTCTGGTATACATCTGGCTATTTAACCAATGAGGAGTTTAAATCAACAGAAACACAGA 2940
Db 2881 TCTCTGGTATACATCTGGCTATTTAACCAATGAGGAGTTTAAATCAACAGAAACACAGA 2940
QY 2941 ATTGATCATCATTTTGCATACCTGCGCATGTAACTCTACTCTGAAATAAATGTTGGTG 3000
Db 2941 ATTGATCATCATTTTGCATACCTGCGCATGTAACTCTACTCTGAAATAAATGTTGGTG 3000
QY 3001 CCATGCGAGGGTCTACGGTTGTGTGTAGTAACTTAATACCTTAACCCCATGCTCAAAA 3060
Db 3001 CCATGCGAGGGTCTACGGTTGTGTGTAGTAACTTAATACCTTAACCCCATGCTCAAAA 3060
QY 3061 TCAATATGATACATATCTCTGAGAGACCCAGCAATACCATAGAAATTAATAAATAAATAA 3120
Db 3061 TCAATATGATACATATCTCTGAGAGACCCAGCAATACCATAGAAATTAATAAATAAATAA 3120

RESULT 3
ABK83327
ID ID
XX ABK83327 standard; cDNA; 4829 BP.
XX
AC ABK83327;
XX
XX
DT 12-AUG-2002 (first entry)
XX
DE cDNA encoding human DPRP-1 splice variant #3.
XX
XX Human; serine protease; dipeptidyl peptidase IV-related protein; DPRP;
KW DPPIV; infection; human immunodeficiency virus; HIV-1; HIV-2; pain;
KW

KW diabetes; infertility; obesity; anorexia; Parkinson's disease; stroke;
KW heart failure; hypertension; urinary retention; osteoporosis; cancer;
KW ulcer; allergy; cancer; psychotic disorder; neurological disorder;
KW dyskinesia; reproductive disorder; inflammatory disorder;
KW metabolic disorder; gene; ss.
OS Homo sapiens.
XX WO200231134-A2.
XX 18-APR-2002.
XX 12-OCT-2001; 2001WO-US031874.
XX 12-OCT-2000; 2000US-0240117P.
XX (FERR) FERRING BV.
XX Qi S, Akinsanya KO, Riviere PJ, Junien J;
XX WPI; 2002-444178/47.
XX P-PSDB; ABG61596.
XX New dipeptidyl peptidase IV-related proteins and nucleic acids encoding
XX the proteins, useful for treating e.g. fungal, bacterial, protozoan and
XX viral infections, cancers, allergies, neurological disorders, or pain.
XX Disclosure; Page 65-66; 113pp; English.
XX The present invention relates to the isolation of novel human serine
XX proteases referred to as dipeptidyl peptidase IV (DPPIV)-related proteins
XX (DPRP). The dipeptidyl peptidase IV-related proteins (DPRP) and nucleic
XX acids encoding them are useful for treating infections such as fungal,
XX bacterial, protozoan and viral infections, particularly infections caused
XX by human immunodeficiency virus (HIV-1 or HIV-2), pain, diabetes,
XX precocious puberty, infertility, obesity, anorexia, bulimia, Parkinson's
XX disease, acute heart failure, hypotension, hypertension, urinary
XX retention, osteoporosis, angina pectoris, stroke, ulcers, asthma,
XX allergies, cancers, migraines, vomiting, psychotic and neurological
XX disorders (e.g. anxiety, dementia, or schizophrenia), and dyskinesias.
XX These may also be used in discovering therapeutic agents for the
XX treatment of reproductive, inflammatory and metabolic disorders. ABK83322
XX -ABK83343 encode human DPRP proteins
SQ Sequence 4829 BP; 1466 A; 886 C; 1017 G; 1460 T; 0 U; 0 Other;
Query Match 99.6%; Score 3106.4; DB 6; Length 4829;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3119; Conservative 0; Mismatches 1; Indels 2; Gaps 1;
QY 1 AAGTGTCTAAAGCTCCGAGGCCAAGCGCGCTGCTACTGCGCGCGCTGCTCTTCTAGTCCG 60
Db 1 AAGTGTCTAAAGCTCCGAGGCCAAGCGCGCTGCTACTGCGCGCGCTGCTCTTCTAGTCCG 60
QY 61 CGTTCCGCGCTGGT 120
Db 61 CGTTCCGCGCTGGT 120
QY 121 GAGTGGAGGCGCGCGAGCATGTAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
Db 121 GAGTGGAGGCGCGCGAGCATGTAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
QY 181 TCCGGGCG 240
Db 181 TCCGGGCG 240
QY 241 CTGGGTGTTGAGATATTTGAAACTGCGGACTGTGAGGAGAAATATTGAATCAAGATCGG 300
Db 241 CTGGGTGTTGAGATATTTGAAACTGCGGACTGTGAGGAGAAATATTGAATCAAGATCGG 300
QY 301 CCTAAATGAGCGCTTTTATGTTAGCGGTATTCTCTGAGTACAGCTTAAAGCTGCTT 360
Db 301 CCTAAATGAGCGCTTTTATGTTAGCGGTATTCTCTGAGTACAGCTTAAAGCTGCTT 360

Qy	361	GCCGATACAGAAAATATCATGCTAGCGCTAAGGCCACCACTGATGTTTCATGTTT	420	1441	ATTGAGTCAGTGCCTGATTCTGTGACGCCACCTAATTATCTATGAAGAAAACAACAGACATC	1500
Db	361		420	1441		1500
Qy	421	GTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGGAATCTATTACCTGCGCAATGCT	480	1501	TGGATAAATATCCATGACATCTTTTCATGTTTTCCTCCCAAAAGTCACGAGAGGAATTTGAG	1560
Db	421		480	1501		1560
Qy	481	GTTGAGAAACAGAGAAAATACACTGTTTATCTCGAAATTCCTGAAATTCCTGAAATCTATCAATAGAGCA	540	1561	TTTATTTTGGCTCTGGAATGCAAAAACAGGTTTCCGTCAATTTATCAAAAATTCATCTATT	1620
Db	481		540	1561		1620
Qy	541	GCAGTCTTAATGCTCTCTTGGAGCCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT	600	1621	TTAAAGGAAAGCAAAATATAAAACGATCCAGTGGTGGCTGCTGCTCCAAAGTATTTCAAG	1680
Db	541		600	1621		1680
Qy	601	GGAAATGTAATCTCGAGAAAGAACTATTAAAGAGAAAGAAAACGCAATGGAAACAGTCGGA	660	1681	TGTCTCTATCAAAAGAGGAGATAGCAATTAACAGTGGTGAATGGGAAGTTCTTGGCCGGCAT	1740
Db	601		660	1681		1740
Qy	661	ATTGCTCTTACGATTTATCAAGGAAGTGGAAACATTTCTGTTTCAAGCCGGTAGTGA	720	1741	GGATCTAATATCCAAAGTTGATGAAATCAGAGGCTGGTATATTTTGAAGGCACCAAGAC	1800
Db	661		720	1741		1800
Qy	721	ATTTATCAGTAAAGATGGAGGCCACAAGGATTTACGCAACACTTTTAAAGGCCAAT	780	1801	TCCCTTTTAGAGCATCAGCTGACGTAGTCACTTACGTAAATCCTGAGAGGTCACAAAGG	1860
Db	721		780	1801		1860
Qy	781	CTAGTGGAAACTAGTGTCTCCCAACATACGAGTGGATCCCAAAATTTAGCCCGCTGATCCA	840	1861	CTGACTGACCGTGGCTACTCACATTTCTGTGTCAGTCAGCACTGTGACTTCTTTATA	1920
Db	781		840	1861		1920
Qy	841	GACTGGATGCTTTTATACATAGCAACGATATTTGGATATCTAACATCGTAAACAGAGAA	900	1921	AGTATGATAGTAAACCAAGGAATTTTGGGCCACCATTTTGGATTCAGCGGT	1980
Db	841		900	1921		1980
Qy	901	GAAAGGAGACTCACTTATGTGCACAATGAGTCAGCCAACTAGGAAGAAGATCCAGATCA	960	1981	GAAGATGACCCAACTTGCAAAAACAAAGGAATTTTGGGCCACCATTTTGGATTCAGCGGT	2040
Db	901		960	1981		2040
Qy	961	GCTGGAGTCGCTACCTTTGTTCTCCAAAGAAAGATTTGATAGATATTTCTGGCTATTGGTGG	1020	2041	CCTCTTCTGACATATCTCTCCAGAAATTTTCTTTTGAAGAGTACTACTGGATTTACA	2100
Db	961		1020	2041		2100
Qy	1021	TGTCNAAAAGCTGAAACAACTCCAGTGGTGGTAAATTTCTTAGAATTTCTATATGAAGAA	1080	2101	TTGTATGGGATGCTCTACAAGCCTCATGATCTACAGCTCGGAAAGAAATATCTTACTGTG	2160
Db	1021		1080	2101		2160
Qy	1081	AATGATGAATCTGAGGTGGAATTTATTTCATGTTACATCCCTATGTTGGAACAAGGAGG	1140	2161	CTGTTTCATATATGTTGG--TCTCTCAGGTGCAAGTTGGTGAATTAATTCGGTTTAAAGGAGTCA	2218
Db	1081		1140	2161		2218
Qy	1141	GCAGATTCATTCGGTTATCTTAACAGGTACAGCAATCTTAAAGTCACTTTTAAAGATG	1200	2219	AGTATTTCCGCTTGTGAATACCTTAGCTCTCTAGGTATGTGTTGTAGTGATAGACAACA	2278
Db	1141		1200	2219		2278
Qy	1201	TCAGAAATTAAGTATGCTGAGGAGGATCATAGATGTCATAGATAGGAATTAATTT	1260	2281	GGGGATCTGTCAACCGGGCTTTAAATTTTGAAGCGCTTTTAAATATAAAAATGGGTCAAA	2338
Db	1201		1260	2281		2338
Qy	1261	CAACCTTTTGAGATCTTAATTTGAAGAGTGTGAATATATTCGAGCTGGATGGACTCT	1320	2339	TAGAAATTAACCATCAGGTGGAGGACTCCCAATATCTAGCTTTCCGATATGATTTTCAATG	2398
Db	1261		1320	2339		2398
Qy	1321	GAGGAAATATGCTTGGTCCATCTAGATCGCTCCAGACTCGCTACAGATAGTG	1380	2341	TAGAAATTAACCATCAGGTGGAGGACTCCCAATATCTAGCTTTCCGATATGATTTTCAATG	2400
Db	1321		1380	2341		2400
Qy	1381	TTGATCTCACTCGAATATTTATCCAGTAGAAGATGATGTTTATGGAAGGCGAGAGCTC	1440	2399	ACTTAGATCGTGGGCAATCCAGCTGGTCTTATGAGGATACCTCTCCCTGATGGCAT	2458
Db	1381		1440	2399		2458
Qy	1440	TTGATCTCACTCGAATATTTATCCAGTAGAAGATGATGTTTATGGAAGGCGAGAGCTC	1440	2401	ACTTAGATCGTGGGCAATCCAGCTGGTCTTATGAGGATACCTCTCCCTGATGGCAT	2460
Db	1440		1440	2401		2460
Qy	1440	TTGATCTCACTCGAATATTTATCCAGTAGAAGATGATGTTTATGGAAGGCGAGAGCTC	1440	2459	TAATGAGAGGTCAGATATCTTCAGGGTGTCTTATGCTGGGCGCCAGTCACTCTGTGGA	2518
Db	1440		1440	2459		2518
Qy	1440	TTGATCTCACTCGAATATTTATCCAGTAGAAGATGATGTTTATGGAAGGCGAGAGCTC	1440	2461	TAATGAGAGGTCAGATATCTTCAGGGTGTCTTATGCTGGGCGCCAGTCACTCTGTGGA	2520
Db	1440		1440	2461		2520
Qy	1440	TTGATCTCACTCGAATATTTATCCAGTAGAAGATGATGTTTATGGAAGGCGAGAGCTC	1440	2519	TCTTCTATGATACAGGATACACGGAAACGTTTATATGGGTCAACCTGACCAAGTAAGAACAGG	2578
Db	1440		1440	2519		2578

QY 431 ATGATCCAGATGGACCTCATTTACAGACAGAAATCTATTACCTTGCCATGTCTGGTGAGAAC 490
Db 446 ATGATCCAGATGGACCTCATTTACAGACAGAAATCTATTACCTTGCCATGTCTGGTGAGAAC 505
QY 491 GAGAAATACACTCTTTTATTCTGAAATCCCAAACTATCAATAGACGACGCTTAA 550
Db 506 GAGAAATACACTCTTTTATTCTGAAATCCCAAACTATCAATAGACGACGCTTAA 565
QY 551 TGCCTCTTGAAGCCTCTTTGGATCTTTTTCAGGCAACACTGGAATGGAATTT 610
Db 566 TGCCTCTTGAAGCCTCTTTGGATCTTTTTCAGGCAACACTGGAATGGAATTT 625
QY 611 CTCGAGAAAGAACTATTAAGAGAAAGAAACGCAATTTGGAACAGTGGGAATTTGCTTT 670
Db 626 CTCGAGAAAGAACTATTAAGAGAAAGAAACGCAATTTGGAACAGTGGGAATTTGCTTT 685
QY 671 ACGATTATCACCAAGGAAGTGGAACTTTCTGTTTCAAGCCGGTAGTGGAAATTTATCAAG 730
Db 686 ACGATTATCACCAAGGAAGTGGAACTTTCTGTTTCAAGCCGGTAGTGGAAATTTATCAAG 745
QY 731 TAAAGATGGAGGGCCCAAGGATTTTACGCAACCAACCTTTAAGGCCCAATCTAGTGGAAA 790
Db 746 TAAAGATGGAGGGCCCAAGGATTTTACGCAACCAACCTTTAAGGCCCAATCTAGTGGAAA 805
QY 791 CTAGTTGTCCCAACATACGATGGATCCAAAATTTATGCCCGCTGATCCAGACTGGATTG 850
Db 806 CTAGTTGTCCCAACATACGATGGATCCAAAATTTATGCCCGCTGATCCAGACTGGATTG 865
QY 851 CTTTATATACATAGCAACGATATTTGGATATCTAAACATCGTAACCAAGAGAAAGGAGAC 910
Db 866 CTTTATATACATAGCAACGATATTTGGATATCTAAACATCGTAACCAAGAGAAAGGAGAC 925
QY 911 TCACCTTAGTGCACAATGAGCTAGCCAAAGTGAAGAAAGTCCAGATCAGCTGGAGTCG 970
Db 926 TCACCTTAGTGCACAATGAGCTAGCCAAAGTGAAGAAAGTCCAGATCAGCTGGAGTCG 985
QY 971 CTAACCTTTGTTCTCAAGAGAAATTTGATAGATATTTCTGGCTATTTGGTGGTGTCCAAAG 1030
Db 986 CTAACCTTTGTTCTCAAGAGAAATTTGATAGATATTTCTGGCTATTTGGTGGTGTCCAAAG 1045
QY 1031 CTGAAACAACTCCAGCTGGTGGTAAATTTCTTAGAATTTCTATAGAAAGAAATGATGAAT 1090
Db 1046 CTGAAACAACTCCAGCTGGTGGTAAATTTCTTAGAATTTCTATAGAAAGAAATGATGAAT 1105
QY 1091 CTGAGGTGGAAATTTATTTATCATTTACATCCCTATGTTGAAAACAAAGGAGGCGAGATTCA 1150
Db 1106 CTGAGGTGGAAATTTATTTATCATTTACATCCCTATGTTGAAAACAAAGGAGGCGAGATTCA 1165
QY 1151 TCCGTTATCTTAAACACAGGTACAGCAAACTCTAAAGTCACTTTTAAAGATGTCAGAAATAA 1210
Db 1166 TCCGTTATCTTAAACACAGGTACAGCAAACTCTAAAGTCACTTTTAAAGATGTCAGAAATAA 1225
QY 1211 TGATTGATGCTGAAGGAAGGATCATAGATGTCATAGATAGGAACTAAATTTCAACCTTTTG 1270
Db 1226 TGATTGATGCTGAAGGAAGGATCATAGATGTCATAGATAGGAACTAAATTTCAACCTTTTG 1285
QY 1271 AGATTCTATTTGAAGGAGTTGAATATATATTTGCCAGAGCTGGATGGACTCTCGAGGGAAT 1330
Db 1286 AGATTCTATTTGAAGGAGTTGAATATATATTTGCCAGAGCTGGATGGACTCTCGAGGGAAT 1345
QY 1331 ATGCTTGGTCCATCTACTAGATCGCTCCAGACTCGCTCAGATAGTGTGATCTCAC 1390
Db 1346 ATGCTTGGTCCATCTACTAGATCGCTCCAGACTCGCTCAGATAGTGTGATCTCAC 1405
QY 1391 CTGAATTTATTTATCCAGTGAAGATGATTTATGAAAGGCGAGACTCATTTAGTGCAG 1450
Db 1406 CTGAATTTATTTATCCAGTGAAGATGATTTATGAAAGGCGAGACTCATTTAGTGCAG 1465
QY 1451 TGCCTGATTTCTGAGCCCACTAAATTTATCTATGAAGAAACAAAGACATCTGGATAATA 1510
Db 1466 TGCCTGATTTCTGAGCCCACTAAATTTATCTATGAAGAAACAAAGACATCTGGATAATA 1525
QY 1511 TCCATGACATCTTTTCATGTTTTCCTCCCAAGTCCAGAGAGGAAATTCAGTTTATTTTGG 1570

Db 1526 TCCATGACATCTTTTCATGTTTTCCTCCCAAGTCCAGAGAGGAAATTCAGTTTATTTTG 1585
QY 1571 CTTCTGAATGCAAAACAGGTTTCGTCATTTATACAAANTTACATCTATTTTAAAGGAAA 1630
Db 1586 CTTCTGAATGCAAAACAGGTTTCGTCATTTATACAAANTTACATCTATTTTAAAGGAAA 1645
QY 1631 GCAATATATAACGATCCAGTGGTGGCTGCTGCTCCAAAGTATTTCAAGTGGCTTATCA 1690
Db 1646 GCAATATATAACGATCCAGTGGTGGCTGCTGCTCCAAAGTATTTCAAGTGGCTTATCA 1705
QY 1691 AAGAGAGATAGCAAAATTAACAGTGGTGAATGGGAAGTTCTTGGCCGCGCATGGAATTA 1750
Db 1706 AAGAGAGATAGCAAAATTAACAGTGGTGAATGGGAAGTTCTTGGCCGCGCATGGAATTA 1765
QY 1751 TCCAAAGTTGATGAAGTACAGAGGCTGGGTATATTTTGAAGGCAACAAAGACTCCCTTTAG 1810
Db 1766 TCCAAAGTTGATGAAGTACAGAGGCTGGGTATATTTTGAAGGCAACAAAGACTCCCTTTAG 1825
QY 1811 AGCATCACTGTGTACGTAGTCAAGTTACGTAAATCTCGAGAGGTGACAAGGCTGACTGACC 1870
Db 1826 AGCATCACTGTGTACGTAGTCAAGTTACGTAAATCTCGAGAGGTGACAAGGCTGACTGACC 1885
QY 1871 GTGGCTACTCAATTTCTGCTGCTCAGTCAAGTCTGCTGCTCTTTTAAAGTAAAGTATA 1930
Db 1886 GTGGCTACTCAATTTCTGCTGCTCAGTCAAGTCTGCTGCTCTTTTAAAGTAAAGTATA 1945
QY 1931 GTAAACAGAGAAATCCACACTGTGTGCTCCTTTTACAAGCTATCAAGTCTCTGAAGATGACC 1990
Db 1946 GTAAACAGAGAAATCCACACTGTGTGCTCCTTTTACAAGCTATCAAGTCTCTGAAGATGACC 2005
QY 1991 CAACTTGCANAAACAAAGAAATTTTGGGCCCAACATTTTGAAGTCTAGTGAATTCATTTGGA 2050
Db 2006 CAACTTGCANAAACAAAGAAATTTTGGGCCCAACATTTTGAAGTCTAGTGAATTCATTTGGA 2065
QY 2051 ACTATACCTCTCCAGAAATTTCTTTTGAAGTACTAGTGAATTCATTTGGAATTCATTTGGA 2110
Db 2066 ACTATACCTCTCCAGAAATTTCTTTTGAAGTACTAGTGAATTCATTTGGAATTCATTTGGA 2125
QY 2111 TGCCTCTACAGCTCATGATCTACAGCTCGAAAGAAATATCTCTAGTGTCTGCTTCAATAT 2170
Db 2126 TGCCTCTACAGCTCATGATCTACAGCTCGAAAGAAATATCTCTAGTGTCTGCTTCAATAT 2185
QY 2171 ATGGTGGTCTCAGGTGAGTTGGTGAATTAATCGGTTTAAAGGAGTCAAGTATTTCCGCT 2230
Db 2186 ATGGTGGTCTCAGGTGAGTTGGTGAATTAATCGGTTTAAAGGAGTCAAGTATTTCCGCT 2245
QY 2231 TGAATACCTTAGCCTCTCTAGGTTATGTGGTTGTAGTGATAGACAACAGGGGATCCTGTG 2290
Db 2246 TGAATACCTTAGCCTCTCTAGGTTATGTGGTTGTAGTGATAGACAACAGGGGATCCTGTG 2305
QY 2291 ACCGAGGGCTTAAATTTTGAAGGCGCTTTTAAATATATAAATGGGTCAAATAGAAATTTGACG 2350
Db 2306 ACCGAGGGCTTAAATTTTGAAGGCGCTTTTAAATATATAAATGGGTCAAATAGAAATTTGACG 2365
QY 2351 ATCAGGTGGAAGGACTCCAAATATCTAGCTTCTCGATATGATTTTCAATTTAGATCGTG 2410
Db 2366 ATCAGGTGGAAGGACTCCAAATATCTAGCTTCTCGATATGATTTTCAATTTAGATCGTG 2425
QY 2411 TGGGCATCCAGGCTGGTCTTATGAGGATACCTCTCCCTGATGGCATTAATCCAGAGT 2470
Db 2426 TGGGCATCCAGGCTGGTCTTATGAGGATACCTCTCCCTGATGGCATTAATCCAGAGT 2485
QY 2471 CAGATATCTTCAGGGTCTTATGCTGGGCGCCAGTCACTCTGTGGAATTCCTATGATA 2530
Db 2486 CAGATATCTTCAGGGTCTTATGCTGGGCGCCAGTCACTCTGTGGAATTCCTATGATA 2545
QY 2531 CAGGATACCGGAAGTTATGAGGTCACTCTGACCAAGATGAACAGGCGCTTACTTAG 2590
Db 2546 CAGGATACCGGAAGTTATGAGGTCACTCTGACCAAGATGAACAGGCGCTTACTTAG 2605
QY 2591 GATCTGTGGCATGCAAGCAAGAAAGTTCCCTCTCGAACCAATTCGTTTACTGCTCTTAC 2650

Db 2606 GATCTGTGGCCATGCAAGCAAGAAAGTTCCCTCTCGAACCAAAATCGTTTACTGCTCTTAC 2665
Qy 2651 ATGGTTTCTCTGATGAGAAATGTCATTTTGCACATACCAAGTATATTTACTGAGTTTTTTAG 2710
Db 2666 ATGGTTTCTCTGATGAGAAATGTCATTTTGCACATACCAAGTATATTTACTGAGTTTTTTAG 2725
Qy 2711 TGAGGGCTGGAAGCCATATGATTTACAGATCTATCTCTAGAGAGACACAGCATAGAG 2770
Db 2726 TGAGGGCTGGAAGCCATATGATTTACAGATCTATCTCTAGAGAGACACAGCATAGAG 2785
Qy 2771 TTCCTGAATCGGAGACATATGATTTACAGATCTATCTCTAGAGAGACACAGCATAGAG 2830
Db 2786 TTCCTGAATCGGAGACATATGATTTACAGATCTATCTCTAGAGAGACACAGCATAGAG 2845
Qy 2831 GATACGATTTGCTGCTCTAAAGCTGATATAATTTTGACCTGTGTAGAACTCTCTGCTAT 2890
Db 2846 GATACGATTTGCTGCTCTAAAGCTGATATAATTTTGACCTGTGTAGAACTCTCTGCTAT 2905
Qy 2891 ACATGGCTATTTAACCAATGAGGAGTTTAAATCAACAGAAACACAGAAATGATCATC 2950
Db 2906 ACATGGCTATTTAACCAATGAGGAGTTTAAATCAACAGAAACACAGAAATGATCATC 2965
Qy 2951 ACATTTGATACCTGCCATGAAATCAATCTACTCTGAAATTAATGTTGGCCATGACGG 3010
Db 2966 ACATTTGATACCTGCCATGAAATCAATCTACTCTGAAATTAATGTTGGCCATGACGG 3025
Qy 3011 GTCACGGTTGTGTAGTATCTAATACCTTAAACCCACATCTCAAAATCAAAATGATA 3070
Db 3026 GTCTACGGTTGTGTAGTATCTAATACCTTAAACCCACATCTCAAAATCAAAATGATA 3085
Qy 3071 CATATTTCTGAGAGACCCAGCAATACCAATAGAAATTAATAAAAAAAAAAAAA 3120
Db 3086 CATATTTCTGAGAGACCCAGCAATACCAATAGAAATTAATAAAAAAAAAAAAA 3135

RESULT 5
ID ABK12892 standard; cDNA; 3106 BP.
XX
AC ABK12892;
XX
DT 09-APR-2002 (first entry)
XX
DE Human protease PRTS-9 cDNA sequence.
XX
KW Human; protease; PRTS; gastrointestinal; Crohn's disease; cancer;
KW cardiovascular; atherosclerosis; autoimmune disorder; dermatitis;
KW inflammatory disorder; acquired immunodeficiency syndrome; AIDS;
KW cell proliferative disorder; developmental disorder; epilepsy;
KW Duchenne muscular dystrophy; epithelial disorder; neurological disorder;
KW reproductive disorder; endometriosis; ss.
XX
OS Homo sapiens.
XX
PH Key Location/Qualifiers
FT CDS 203..2851
FT /*tag= a
FT /product= "Human protease PRTS-9"
XX
WO200198468-A2.
XX
PD 27-DEC-2001.
XX
PF 13-JUN-2001; 2001WO-US019178.
XX
PR 16-JUN-2000; 2000US-0212336P.
PR 22-JUN-2000; 2000US-0213955P.
PR 29-JUN-2000; 2000US-0215396P.
PR 07-JUL-2000; 2000US-0216821P.
PR 14-JUL-2000; 2000US-0218946P.
XX
PA (INCYTE) INCYTE GENOMICS INC.
XX

PI Yue H, Elliott VS, Gandhi AR, Lal P, Au-Young J, Tribouley CW,
PI Deleane AM, Baughn MR, Nguyen DB, Lee EA, Hafalia A, Khan FA,
PI Wallia NK, Yao MG, Lu DM, Patterson C, Tang YT, Walsh RT;
PI Azimzai Y, Lu Y, Ramkumar J, Xu Y, Reddy R, Das D, Kearney L;
PI Kallick DA;
XX
DR WPI; 2002-090437/12.
DR P-PSDB; AAU74749.
XX
PT Twenty one human proteases (referred to as PRTS-1 to PRTS-21), useful in
PT the diagnosis, treatment and prevention of gastrointestinal (e.g.
PT gastritis), cardiovascular (e.g. atherosclerosis) and cell proliferative
PT (e.g. cancer) disorders.
XX
PS Claim 5; Page 166-167; 177pp; English.
XX
CC The present invention relates to twenty one new human proteases, referred
CC to as PRTS-1 to PRTS-21. The PRTS polynucleotides and polypeptides of the
CC invention are useful in the diagnosis, treatment and prevention of
CC gastrointestinal e.g. gastritis, esophageal carcinoma and Crohn's
CC disease, cardiovascular e.g. atherosclerosis, hypertension and myocardial
CC infarction, autoimmune/inflammatory e.g. acquired immunodeficiency
CC syndrome (AIDS), allergies and osteoarthritis, cell proliferative e.g.
CC cancer, developmental e.g. Duchenne and Becker muscular dystrophy,
CC epithelial e.g. dermatitis, neurological e.g. epilepsy and Alzheimer's
CC disease and reproductive e.g. infertility and endometriosis disorders.
CC Numerous other examples of each disorder are given in the specification.
CC The present nucleic acid sequence encodes the human protease PRTS-9
CC protein of the invention
XX
SQ Sequence 3106 BP; 928 A; 633 C; 704 G; 841 T; 0 U; 0 Other;

Query Match 99.4%; Score 3101.8; DB 6; Length 3106;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3103; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 13 CTCGAGGCCAAGGCGCGCTGCTACTGCGCGCGCTGCTTCTTAGTCCGCGCTTCCGCGCCT 72
Db 2 CTCGAGGCCAAGGCGCGCTGCTACTGCGCGCGCTGCTTCTTAGTCCGCGCTTCCGCGCCT 61

Qy 73 GGGTTCTACCGG 132
Db 62 GGGTTCTACCGG 121

Qy 133 CGCAGCATGAAGCGG 192
Db 122 CGCAGCATGAAGCGG 181

Qy 193 CGGGGGGAAGGAAATGCAACATGCGAGCGAGCAATGGAACACAGACAGCTGGTGTGAG 252
Db 182 CGGGGGGAAGGAAATGCAACATGCGAGCGAGCAATGGAACACAGACAGCTGGTGTGAG 241

Qy 253 ATATTTGAAACTGCGGACTGTGAGGAGATATTTGAATCAGAGGATCGGCTAAATTTGAG 312
Db 242 ATATTTGAAACTGCGGACTGTGAGGAGATATTTGAATCAGAGGATCGGCTAAATTTGAG 301

Qy 313 CCTTTTATGTTGAGCGGATTTCTTGGAGTCAAGCTTAAAGAGCTCTTCCGCGATACCA 372
Db 302 CCTTTTATGTTGAGCGGATTTCTTGGAGTCAAGCTTAAAGAGCTCTTCCGCGATACCA 361

Qy 373 AAATATCATGCGCTACATGATGCTAAGGACCAACATGATTTTGTGAGAGGAGAT 432
Db 362 AAATATCATGCGCTACATGATGCTAAGGACCAACATGATTTTGTGAGAGGAGAT 421

Qy 433 GATCCAGATGAGCTCATTCAGACAGATCTTATACCTTGCCATGCTGCTGAGACAGAG 492
Db 422 GATCCAGATGAGCTCATTCAGACAGATCTTATACCTTGCCATGCTGCTGAGACAGAG 481

Qy 493 GAAATATACATGTTTATTTCTGAAATTTCCCAAACTATCAATAGACAGAGCTCTTAATG 552
Db 482 GAAATATACATGTTTATTTCTGAAATTTCCCAAACTATCAATAGACAGAGCTCTTAATG 541

Qy 553 CTCTCTTGGAGGCTCTTTTGGATCTTTTTCAGGCAACACTGAGATATGGAATGTATTCT 612

Db 542 |||||CTCTCTGGAGGCTCTTTTGGATCTTTTTCAGGCAACACTGGACTATGGAAATGTAATCT||| 601
Qy 613 CGAGAAAGAACTATTAAGAGAAAGAAACGCAATGGAAACAGTCGGAAATGCTTCTTAC 672
Db 602 CGAGAAAGAACTATTAAGAGAAAGAAACGCAATGGAAACAGTCGGAAATGCTTCTTAC 661
Qy 673 GATTATCAACGAAGAGTGGAAACATTTCTCTTTCAAAGCCGGTAGTGGAAATTTATCACGTA 732
Db 662 GATTATCAACGAAGAGTGGAAACATTTCTCTTTCAAAGCCGGTAGTGGAAATTTATCACGTA 721
Qy 733 AAAAGTGAAGGGCCACAAAGGATTTACGCAACAACCTTTAAGGCCCAATCTAGTGGAAACT 792
Db 722 AAAGATGAAGGGCCACAAAGGATTTACGCAACAACCTTTAAGGCCCAATCTAGTGGAAACT 781
Qy 793 AGTTGTCCTCAACATACGGATCGGATCCGCAATTTATGCCCGCTGATCCAGACTGGATTCCT 852
Db 782 AGTTGTCCTCAACATACGGATCGGATCCGCAATTTATGCCCGCTGATCCAGACTGGATTCCT 841
Qy 853 TTTATACATAGCAACGATATTTGGATATCTAAACATCGTAAACAGAGAAAGAGAGACTC 912
Db 842 TTTATACATAGCAACGATATTTGGATATCTAAACATCGTAAACAGAGAAAGAGAGACTC 901
Qy 913 ACTTATGTGCACATGAGCTAGCCAAACATCGAAGAGATGCCAGATCAGCTCGAGTCGCT 972
Db 902 ACTTATGTGCACATGAGCTAGCCAAACATCGAAGAGATGCCAGATCAGCTCGAGTCGCT 961
Qy 973 ACCTTTGTTCTCCAAGAAAGATTTGATAGATATTTCTGGCTATTTGGTGGTGTCCAAAGCT 1032
Db 962 ACCTTTGTTCTCCAAGAAAGATTTGATAGATATTTCTGGCTATTTGGTGGTGTCCAAAGCT 1021
Qy 1033 GAAACAACTCCAGTGGTGGTAAATTTCTTAGAAATTTCTATATGAAGAAATGATGAATCT 1092
Db 1022 GAAACAACTCCAGTGGTGGTAAATTTCTTAGAAATTTCTATATGAAGAAATGATGAATCT 1081
Qy 1093 GAGGTGGAATTTATTCATGTTTACATCCCTATGTTGGAAACAGAGAGGCGAGATTCATTC 1152
Db 1082 GAGGTGGAATTTATTCATGTTTACATCCCTATGTTGGAAACAGAGAGGCGAGATTCATTC 1141
Qy 1153 CGTTATCTTAAACAGGTACAGCAAACTCTTAAAGTCACTTTTAAAGTGCAGAAATAATG 1212
Db 1142 CGTTATCTTAAACAGGTACAGCAAACTCTTAAAGTCACTTTTAAAGTGCAGAAATAATG 1201
Qy 1213 ATTGATGCTGAAGGAGGATCATAGATGTCATAGATAAGGAACCTAATTCACCTTTTGAG 1272
Db 1202 ATTGATGCTGAAGGAGGATCATAGATGTCATAGATAAGGAACCTAATTCACCTTTTGAG 1261
Qy 1273 ATTCCTATTGAAGGAGTTGAATATATTGCCAGAGCTGGATGGACTCTCGAGGGAATAAT 1332
Db 1262 ATTCCTATTGAAGGAGTTGAATATATTGCCAGAGCTGGATGGACTCTCGAGGGAATAAT 1321
Qy 1333 GCTTGGTCATCTCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCAGCT 1392
Db 1322 GCTTGGTCATCTCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCAGCT 1381
Qy 1393 GAAATTTATTTATCCAGTAGAAGATGATGTTATGGAAAGGACAGACTCAATGAGTCAGTG 1452
Db 1382 GAAATTTATTTATCCAGTAGAAGATGATGTTATGGAAAGGACAGACTCAATGAGTCAGTG 1441
Qy 1453 CTTGATCTGTGAAGCCCACTAATTTATCTATGAAGAAACAAACAGACACTGTTGAATAATATC 1512
Db 1442 CTTGATCTGTGAAGCCCACTAATTTATCTATGAAGAAACAAACAGACACTGTTGAATAATATC 1501
Qy 1513 CATGACATCTTTTATGTTTTTCCCAAAGTCAAGAGGAAATTCAGTGTATTTTGTGCC 1572
Db 1502 CATGACATCTTTTATGTTTTTCCCAAAGTCAAGAGGAAATTCAGTGTATTTTGTGCC 1561
Qy 1573 TCTGAAATGCAAAACAGGTTTCCCGTCATTTTATACAAATATCATCTATTTTAAAGGAAAGC 1632
Db 1562 TCTGAAATGCAAAACAGGTTTCCCGTCATTTTATACAAATATCATCTATTTTAAAGGAAAGC 1621
Qy 1633 AAATATAACCATCCAGTGGTGGCTGCCTGCTCCAGTGATTTCAAGTGTCTCTATCAAA 1692

Db 1622 AAATATAACCATCCAGTGGTGGCTGCTGCTCCAAGTGATTTCAAGTGTCTCTATCAAA 1681
Qy 1693 GAGGAGATAGCAAAATTTACCAAGTGGTGAATGGGAAGTCTTCTGGCCGGATGGATCTAATATC 1752
Db 1682 GAGGAGATAGCAAAATTTACCAAGTGGTGAATGGGAAGTCTTCTGGCCGGCATGGATCTAATATC 1741
Qy 1753 CAAAGTTGATGAAGTCAAGAGGCTGGTATATTTTGAAGCCACCAAGACTCCCTCTTTAGAG 1812
Db 1742 CAAAGTTGATGAAGTCAAGAGGCTGGTATATTTTGAAGCCACCAAGACTCCCTCTTTAGAG 1801
Qy 1813 CATCACTGTATAGTACGTATTCAGTTACGTAAATCTCTGAGAGGTGACAAAGCTGACACCGT 1872
Db 1802 CATCACTGTATAGTACGTATTCAGTTAAATCTCTGAGAGGTGACAAAGCTGACACCGT 1861
Qy 1873 GGCTACTCACTTCCTGCTGCTCAGTCAGCACTGTGACTTCTTTATAAGTAAGTATAGT 1932
Db 1862 GGCTACTCACTTCCTGCTGCTCAGTCAGCACTGTGACTTCTTTATAAGTAAGTATAGT 1921
Qy 1933 AACCAAGAAATCCACACTGTGTGCTCCCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCA 1992
Db 1922 AACCAAGAAATCCACACTGTGTGCTCCCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCA 1981
Qy 1993 ACTTGCAAAACAAAGGAAATTTTGGGCCACCACTTTTGGATTTCAGCAGGTCTCTTCTGTAC 2052
Db 1982 ACTTGCAAAACAAAGGAAATTTTGGGCCACCACTTTTGGATTTCAGCAGGTCTCTTCTGTAC 2041
Qy 2053 TATATCTCTCCAGAAATTTCTCTTTTGAAGTACTACTGGATTTTACATTTGATGGATG 2112
Db 2042 TATATCTCTCCAGAAATTTCTCTTTTGAAGTACTACTGGATTTTACATTTGATGGATG 2101
Qy 2113 CTCCTACAGCCTCATGATCTACAGCTTGGAAAGAAATATCTCTACTGTGCTGCTCATATAT 2172
Db 2102 CTCCTACAGCCTCATGATCTACAGCTTGGAAAGAAATATCTCTACTGTGCTGCTCATATAT 2161
Qy 2173 GGTGGTCTCTCAGGTGCAGTTCGGTGAAATATCGGTTTAAAGAGTCAAGTATTTCCGCTTG 2232
Db 2162 GGTGGTCTCTCAGGTGCAGTTCGGTGAAATATCGGTTTAAAGAGTCAAGTATTTCCGCTTG 2221
Qy 2233 AATACTCAGCTCTCTAGGTATATGTTGGTTGTAGTATAGACAAACAGGGGATCTCTGTAC 2292
Db 2222 AATACTCAGCTCTCTAGGTATATGTTGGTTGTAGTATAGACAAACAGGGGATCTCTGTAC 2281
Qy 2293 CGAGGCTCTTAAATTTGAAGGGCGCTTTAAATATATAAATGGGTCAAAATAGAAAATGACGAT 2352
Db 2282 CGAGGCTCTTAAATTTGAAGGGCGCTTTAAATATATAAATGGGTCAAAATAGAAAATGACGAT 2341
Qy 2353 CAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTCAATTTGACTTTAGATCGTGTG 2412
Db 2342 CAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTCAATTTGACTTTAGATCGTGTG 2401
Qy 2413 GGCATCCAGCGCTGGTCTTATGGAAGTATACCTCTCCCTGATGGCATTTAATGCAAGGTCA 2472
Db 2402 GGCATCCAGCGCTGGTCTTATGGAAGTATACCTCTCCCTGATGGCATTTAATGCAAGGTCA 2461
Qy 2473 GATATCTTCAGGGTTCATTTGCTGGGGCCCGAGTCACTCTGTGGATCTTCTATGATACA 2532
Db 2462 GATATCTTCAGGGTTCATTTGCTGGGGCCCGAGTCACTCTGTGGATCTTCTATGATACA 2521
Qy 2533 GGNATACAGGAACGTTATATGGGTCACTCTGACCAAGATGAAACAGGGCTATTTAGGA 2592
Db 2522 GGNATACAGGAACGTTATATGGGTCACTCTGACCAAGATGAAACAGGGCTATTTAGGA 2581
Qy 2593 TCTGTGGCCATCCAAAGCAGAAAAGTTCCTCTGAAACAAATCGTTTACTGTCTTTACAT 2652
Db 2582 TCTGTGGCCATCCAAAGCAGAAAAGTTCCTCTGAAACAAATCGTTTACTGTCTTTACAT 2641
Qy 2653 GGTTCCTCGATGAGAAATGTCATTTTGCACATACCAGTATTTACTGTAGTTTTTTAGTG 2712
Db 2642 GGTTCCTCGATGAGAAATGTCATTTTGCACATACCAGTATTTACTGTAGTTTTTTAGTG 2701
Qy 2713 AGGGCTGGAAGGCCATATGATTTTACAGATCTATCTCTGAGAGACACAGCATAAAGATT 2772
Db 2702 AGGGCTGGAAGGCCATATGATTTTACAGATCTATCTCTCAGGAGACACAGCATAAAGATT 2761

QY 2773 CCTGAATCGGAGAACATTAATGAACTGCACTCTTTTGCACTACCTTCAAGAAAACCTTGA 2832
DB 2762 CCTGAATCGGAGAACATTAATGAACTGCACTCTTTTGCACTACCTTCAAGAAAACCTTGA 2821
QY 2833 TCACGTATGCTGCTCTAAAGTGATATAATTTTACCTGTGTAGAACTCTCTGGTATAC 2892
DB 2822 TCACGTATGCTGCTCTAAAGTGATATAATTTTACCTGTGTAGAACTCTCTGGTATAC 2881
QY 2893 ACTGGCTATTTAAACCAATGAGGAGGTTTAAATCAACAGAAAACACAGAAATTCATCATAC 2952
DB 2882 ACTGGCTATTTAAACCAATGAGGAGGTTTAAATCAACAGAAAACACAGAAATTCATCATAC 2941
QY 2953 ATTTTGATACCTGCGCATGTAACATCTACTCTGAAATAAATGTGGTGCATCGAGGGT 3012
DB 2942 ATTTTGATACCTGCGCATGTAACATCTACTCTGAAATAAATGTGGTGCATCGAGGGT 3001
QY 3013 CTACGGTTTGGTAGTAATCTTAATACCTTAAACCCACATGCTCAAAATCAATGATACA 3072
DB 3002 CTACGGTTTGGTAGTAATCTTAATACCTTAAACCCACATGCTCAAAATCAATGATACA 3061
QY 3073 TATTCCTGAGAGACCCAGCAATACATAGAAATTTACTAAAAAAA 3117
DB 3062 TATTCCTGAGAGACCCAGCAATACATAGAAATTTACTAAAAAAA 3106

RESULT 6

ABK83332
ID ABK83332 standard; cDNA; 4685 BP.

AC ABK83332;

DT 12-AUG-2002 (first entry)

DE cDNA encoding human DPRP-1 splice variant #8.

XX Human; serine protease; dipeptidyl peptidase IV-related protein; DPRP;
XX DPPIV; infection; human immunodeficiency virus; HIV-1; HIV-2; pain;
XX diabetes; infertility; obesity; anorexia; Parkinson's disease; stroke;
XX heart failure; hypertension; urinary retention; osteoporosis; cancer;
XX ulcer; allergy; cancer; psychotic disorder; neurological disorder;
XX dyskinesia; reproductive disorder; inflammatory disorder;
XX metabolic disorder; gene; ss.

OS Homo sapiens.

PN WO200231134-A2.

PD 18-APR-2002.

XX 12-OCT-2001; 2001WO-US031874.

XX 12-OCT-2000; 2000US-0240117P.

XX (FERR) FERRING BV.

PA Qi S, Akinsanya KO, Riviere PJ, Junien J;

XX WPI: 2002-444178/47.

XX P-PSDB; ABG61601.

XX New dipeptidyl peptidase IV-related proteins and nucleic acids encoding
PT the proteins, useful for treating e.g. fungal, bacterial, protozoan and
PT viral infections, cancers, allergies, neurological disorders, or pain.

XX Disclosure; Page 75-76; 113pp; English.

XX The present invention relates to the isolation of novel human serine
CC proteases referred to as dipeptidyl peptidase IV (DPPIV)-related proteins
CC (DPRP). The dipeptidyl peptidase IV-related proteins (DPRP) and nucleic
CC acids encoding them are useful for treating infections such as fungal,
CC bacterial, protozoan and viral infections, particularly infections caused
CC by human immunodeficiency virus (HIV-1 or HIV-2), pain, diabetes,

CC precocious puberty, infertility, obesity, anorexia, bulimia, Parkinson's
CC disease, acute heart failure, hypotension, hypertension, urinary
CC retention, osteoporosis, angina pectoris, stroke, ulcers, asthma,
CC allergies, cancers, migraine, vomiting, psychotic and neurological
CC disorders (e.g. anxiety, dementia, or schizophrenia), and dyskinesias.
CC These may also be used in discovering therapeutic agents for the
CC treatment of reproductive, inflammatory and metabolic disorders. ABK83322
CC -ABK83343 encode human DPRP proteins

XX SQ Sequence 4685 BP; 1430 A; 853 C; 991 G; 1411 T; 0 U; 0 Other;

Query Match 90.5%; Score 2824.4; DB 6; Length 4685;
Best Local Similarity 95.4%; Pred. No. 0;
Matches 2977; Conservative 0; Mismatches 1; Indels 142; Gaps 1;

QY 1 AAGTGTCTAAGCCTCCGAGGCCAAGCCGCTGCTACTGCCGCCGCTGCTTTTATGTCGG 60
DB 1 AAGTGTCTAAGCCTCCGAGGCCAAGCCGCTGCTACTGCCGCCGCTGCTTTTATGTCGG 60
QY 61 CGTTCCGCCCTGGTGTGTACCGCGCGCGCGCGAGGAAGCCACTGCAACACGAGACCG 120
DB 61 CGTTCCGCCCTGGTGTGTACCGCGCGCGCGCGAGGAAGCCACTGCAACACGAGACCG 120
QY 121 GAGTGGAGGGCGGCGAGCATGAAGCGCGCGCGCTCCATAGCGCACTCGGGACGG 180
DB 121 GAGTGGAGGGCGGCGAGCATGAAGCGCGCGCGCTCCATAGCGCACTCGGGACGG 180
QY 181 TCCGGCGCGCGCGGGGGAAGAAAATGCAACATGGCAGCAGCAATGGAACAGACAG 240
DB 181 TCCGGCGCGCGCGGGGGAAGAAAATGCAACATGGCAGCAGCAATGGAACAGACAG 240
QY 241 CTGGGTGTTGAGATATTTGAAACTCGGACCTGTGAGGAGATATTTGAATCAAGATCGG 300
DB 241 CTGGGTGTTGAGATATTTGAAACTCGGACCTGTGAGGAGATATTTGAATCAAGATCGG 300
QY 301 CCTAAATTTGAGCCTTTTATGTTGAGCGGTATTTCTGAGCTCAGCTTAAAGCTGCTT 360
DB 301 CCTAAATTTGAGCCTTTTATGTTGAGCGGTATTTCTGAGCTCAGCTTAAAGCTGCTT 360
QY 361 GCCGATACCAAGAAAATATCATGCTACATGATGGCTTAAGGCCACCACATGATTTATGTTT 420
DB 361 GCCGATACCAAGAAAATATCATGCTACATGATGGCTTAAGGCCACCACATGATTTATGTTT 420
QY 421 GTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGAACTATTACTCTTGCCATGCT 480
DB 421 GTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGAACTATTACTCTTGCCATGCT 480
QY 481 GGTGAGAACAGAGAAAATACACTGTTTATTTCTGAAATTTCCCAAACTATCAATAGACGA 540
DB 481 GGTGAGAACAGAGAAAATACACTGTTTATTTCTGAAATTTCCCAAACTATCAATAGACGA 540
QY 541 GCAGTCTTAATGCTCTCTTGGAGGCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT 600
DB 541 GCAGTCTTAATGCTCTCTTGGAGGCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT 600
QY 601 GGAATGTATTCTCGAGAAGAACTATTATTAAGAGAAAAGAAAACGCAATTTGGAACAGTCGA 660
DB 601 GGAATGTATTCTCGAGAAGAACTATTATTAAGAGAAAAGAAAACGCAATTTGGAACAGTCGA 660
QY 661 ATTGCTTTTACGATTATCACCAAGGAAGTGGAACTTTCTGTTTCAAGCCGGTAGTGA 720
DB 661 ATTGCTTTTACGATTATCACCAAGGAAGTGGAACTTTCTGTTTCAAGCCGGTAGTGA 720
QY 721 ATTTATCACGTAAAGTGGAGGCGGCACAGGATTTAGCACAACCTTTAAGGCCCAAT 780
DB 721 ATTTATCACGTAAAGTGGAGGCGGCACAGGATTTAGCACAACCTTTAAGGCCCAAT 780
QY 781 CTAGTGGAAACTAGTGTGTTCCCAACATACGATGGATCCAAATTTATGCTGCTATCA 840
DB 781 CTAGTGGAAACTAGTGTGTTCCCAACATACGATGGATCCAAATTTATGCTGCTATCA 840
QY 841 GACTGGATTGCTTTTATACATACAGATATTTGGATATCTAAATCGTAAACAGAGAA 900
DB 841 GACTGGATTGCTTTTATACATACAGATATTTGGATATCTAAATCGTAAACAGAGAA 900

Db	841	GACTGATTGCTTTTATACATAGCAACGATATTTGGATATCTAAACATCGTAACACAGAGAA	900
Qy	901	GAAGGAGACTCACATTATGTCGCAAAATGAGCTAGCCAAACATGGAAGACATGCCAGATCA	960
Db	901	GAAGGAGACTCACATTATGTCGCAAAATGAGCTAGCCAAACATGGAAGACATGCCAGATCA	960
Qy	961	GCTGAGTCCGCTACCTTTGTTCTCCAAAGAGAAATTTGATAGATATTTCTGGCTATTTGGTG	1020
Db	961	GCTGAGTCCGCTACCTTTGTTCTCCAAAGAGAAATTTGATAGATATTTCTGGCTATTTGGTG	1020
Qy	1021	TGTCGAAAGCTGAAACAACTCCGAGTGGTGGTAAATTTCTTAGAATTTCTATATGAAGAA	1080
Db	1021	TGTCGAAAGCTGAAACAACTCCGAGTGGTGGTAAATTTCTTAGAATTTCTATATGAAGAA	1080
Qy	1081	AATGATGAATCTGAGGTGGAATTAATTCATGTTACATCCCTATGTTGGAACAAAGGAGG	1140
Db	1081	AATGATGAATCTGAGGTGGAATTAATTCATGTTACATCCCTATGTTGGAACAAAGGAGG	1140
Qy	1141	GCAGATTCATTCGCTTATCCCTAAACAGGTACAGCAAACTCCTAAAGTCACTTTTAAGATG	1200
Db	1141	GCAGATTCATTCGCTTATCCCTAAACAGGTACAGCAAACTCCTAAAGTCACTTTTAAGATG	1200
Qy	1201	TCAGAAATTAATGATGCTGGAAGGAGGATCATAGATGTCATAGATAAGGAATCAATTT	1260
Db	1201	TCAGAAATTAATGATGCTGGAAGGAGGATCATAGATGTCATAGATAAGGAATCAATTT	1260
Qy	1261	CAACCTTTTGAGATTTCTATTTGAAGGAGTTGAAATATATTGCCAGAGCTGGATGACTCT	1320
Db	1261	CAACCTTTTGAGATTTCTATTTGAAGGAGTTGAAATATATTGCCAGAGCTGGATGACTCT	1320
Qy	1321	GAGGAAATATGCTTGGTCCATCTACTAGATCGCTCCGAGCTCGCTACAGATAGTG	1380
Db	1321	GAGGAAATATGCTTGGTCCATCTACTAGATCGCTCCGAGCTCGCTACAGATAGTG	1380
Qy	1381	TTGATCTCACCTGAAATATTTTATCCGAGTAGAGATGATGTTATGAAAGGACAGACTC	1440
Db	1381	TTGATCTCACCTGAAATATTTTATCCGAGTAGAGATGATGTTATGAAAGGACAGACTC	1440
Qy	1441	ATTGAGTCAGTGCCTGATTCGTGACGCCACTAATTAATCTATGAAGAAACAAACAGACATC	1500
Db	1441	ATTGAGTCAGTGCCTGATTCGTGACGCCACTAATTAATCTATGAAGAAACAAACAGACATC	1500
Qy	1501	TGGATAAATATCCATGACATCTTTTATGTTTTTCCCAAGTCAACGAGAGAAATTTGAG	1560
Db	1501	TGGATAAATATCCATGACATCTTTTATGTTTTTCCCAAGTCAACGAGAGAAATTTGAG	1560
Qy	1561	TTTATTTTGGCTCTGAAATGCAAAACAGGTTTCCGTCATTTATACAAAATTTACATCTATT	1620
Db	1561	TTTATTTTGGCTCTGAAATGCAAAACAGGTTTCCGTCATTTATACAAAATTTACATCTATT	1620
Qy	1621	TTAAAGGAAAGCAAAATATAACGATCCAGTGGTGGCTGCTCCTCCAAAGTCAATTTCAAG	1680
Db	1621	TTAAAGGAAAGCAAAATATAACGATCCAGTGGTGGCTGCTCCTCCAAAGTCAATTTCAAG	1680
Qy	1681	TGTCCTATCAAGGAGATAGCAATTAACAGTGGTGAATGGAGGTTCTTGGCCGGCAT	1740
Db	1681	TGTCCTATCAAGGAGATAGCAATTAACAGTGGTGAATGGAGGTTCTTGGCCGGCAT	1740
Qy	1741	GGATCTAATATCCAAAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCCACCAAGAC	1800
Db	1741	GGATCTAATATCCAAAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCCACCAAGAC	1800
Qy	1801	TCCCTTTTAGAGCATACCTGTACGTAGTCAGTTACGTAAATCCTGGAGAGGTGACAAAGG	1860
Db	1801	TCCCTTTTAGAGCATACCTGTACGTAGTCAGTTACGTAAATCCTGGAGAGGTGACAAAGG	1860
Qy	1861	CTGACTGACCGTGGCTACTCAATTTCTGCTGCAATGTCAGCACTGTCACTTCTTTATA	1920
Db	1861	CTGACTGACCGTGGCTACTCAATTTCTGCTGCAATGTCAGCACTGTCACTTCTTTATA	1920
Qy	1921	AGTAAGTAGTAGTAACCAAGAAATCCCACTGTGTGTCCTTTTACAAAGCTATCAAGTCTCT	1980
Db	1921	AGTAAGTAGTAGTAACCAAGAAATCCCACTGTGTGTCCTTTTACAAAGCTATCAAGTCTCT	1980

Qy	1981	GAAGATGACCCAACTTTGCAAAACAAAGGAAATTTTGGGCCACCAATTTTGGATTTGACGAGGT	2040
Db	1981	GAAGATGACCCAACTTTGCAAAACAAAGGAAATTTTGGGCCACCAATTTTGGATTTCA	2034
Qy	2041	CCTCTCTCTGACTATACTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTTACA	2100
Db	2035	-----	2034
Qy	2101	TTGTATGGGATGCTCTACAGCCTCATGATCTACAGCCTGGAAAGAAATATCTCTACTGTG	2160
Db	2035	-----	2034
Qy	2161	CTGTTCAATATATGGTGGTCTCAGGTGAGTGTGGTGAATTAATCGGTTTAAAGGAGTCAAG	2220
Db	2035	-----GTCCTCAGGTGAGTGTGGTGAATTAATCGGTTTAAAGGAGTCAAG	2078
Qy	2221	TATTTCCGCTTGAATACCTTAGCCTCTCTAGGTTATGTGGTTGTAGTAGACAACAGG	2280
Db	2079	TATTTCCGCTTGAATACCTTAGCCTCTCTAGGTTATGTGGTTGTAGTAGACAACAGG	2138
Qy	2281	GGATCCTGTCAACCGAGGCTTAAATTTGAAGCGCCTTTAAATATAAAATGGGTCAAATA	2340
Db	2139	GGATCCTGTCAACCGAGGCTTAAATTTGAAGCGCCTTTAAATATAAAATGGGTCAAATA	2198
Qy	2341	GAAATTTGACGATCAGGTGGAAGGACTCCAAATATCTAGCTTCTCGATATGATTTCAATTGAC	2400
Db	2199	GAAATTTGACGATCAGGTGGAAGGACTCCAAATATCTAGCTTCTCGATATGATTTCAATTGAC	2258
Qy	2401	TTAGATCCTGTGGGCATCCAGGCTGGTCTTATGGAGGATACCTCTCCCTGATGGCATTTA	2460
Db	2259	TTAGATCCTGTGGGCATCCAGGCTGGTCTTATGGAGGATACCTCTCCCTGATGGCATTTA	2318
Qy	2461	ATGCAGAGTCTAGATATCTTCAGGGTTCCTAGCTGGGCCCGAGTCACTCTGTGGATC	2520
Db	2319	ATGCAGAGTCTAGATATCTTCAGGGTTCCTAGCTGGGCCCGAGTCACTCTGTGGATC	2378
Qy	2521	TTCTATGATACAGGATACAACGGAAGTTATATGGGTCACTCTGACACAGAAATGAACAGGGC	2580
Db	2379	TTCTATGATACAGGATACAACGGAAGTTATATGGGTCACTCTGACACAGAAATGAACAGGGC	2438
Qy	2581	TATTAATTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGAAACCAAAATCGTTTA	2640
Db	2439	TATTAATTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGAAACCAAAATCGTTTA	2498
Qy	2641	CTGCTCTTACATGGTTTCTCGATGAGAAATGTCATTTTGACATACACAGTATATTTACTG	2700
Db	2499	CTGCTCTTACATGGTTTCTCGATGAGAAATGTCATTTTGACATACACAGTATATTTACTG	2558
Qy	2701	AGTTTTTTAGTCAGGGCTGGAAAGCCATATGATTTTACAGATCTATCTCTCAGGAGACAC	2760
Db	2559	AGTTTTTTAGTCAGGGCTGGAAAGCCATATGATTTTACAGATCTATCTCTCAGGAGACAC	2618
Qy	2761	AGCATAGAGTTTCTTGAATCGGAGAACATTAATGAATGCACTGCACTTTTGGCACTACCTTCAA	2820
Db	2619	AGCATAGAGTTTCTTGAATCGGAGAACATTAATGAATGCACTGCACTTTTGGCACTACCTTCAA	2678
Qy	2821	GAAGACCTTGGATCAGCTATTTGCTCTTAAAGTGATATATTTTTCACCTGTGTAGAAC	2880
Db	2679	GAAGACCTTGGATCAGCTATTTGCTCTCTTAAAGTGATATATTTTTCACCTGTGTAGAAC	2738
Qy	2881	TCTCTGGTATACATCGGCTATTTTAAACAAATGAGAGGTTTAAATCAACAGAAACACAGA	2940
Db	2739	TCTCTGGTATACATCGGCTATTTTAAACAAATGAGAGGTTTAAATCAACAGAAACACAGA	2798
Qy	2941	ATTGATCATCAATTTTGATACCTGGCCATGTAACATCTCTCTGAAAATAAATGTGGTG	3000
Db	2799	ATTGATCATCAATTTTGATACCTGGCCATGTAACATCTCTCTGAAAATAAATGTGGTG	2858
Qy	3001	CCAATGACGGGTCTACGGTTTGTGGTAGTAATCTAAATACCTTAAACCCCAATGCTCAAAA	3060
Db	2859	CCAATGACGGGTCTACGGTTTGTGGTAGTAATCTAAATACCTTAAACCCCAATGCTCAAAA	2918

QY 3061 TCAATGATACATATCTCTGAGAGACCCAGCAATACCATAGAAATTTACTTAAAAA 3120
Db 2919 TCAATGATACATATCTCTGAGAGACCCAGCAATACCATAGAAATTTACTTAAAAA 2978

RESULT 7

ID ABK83331 standard; cDNA; 4676 BP.

AC ABK83331;

DT 12-AUG-2002 (first entry)

DE cDNA encoding human DPRP-1 splice variant #7.

XX Human; serine protease; dipeptidyl peptidase IV-related protein; DPRP;
KW DPPIV; infection; human immunodeficiency virus; HIV-1; HIV-2; pain;
KW diabetes; infertility; obesity; anorexia; Parkinson's disease; stroke;
KW heart failure; hypertension; urinary retention; osteoporosis; cancer;
KW ulcer; allergy; cancer; psychotic disorder; neurological disorder;
KW dyskinnesia; reproductive disorder; inflammatory disorder;
KW metabolic disorder; gene; ss.

XX Homo sapiens.

OS WO200231134-A2.

PN 18-APR-2002.

XX 12-OCT-2001; 2001WO-US031874.

XX 12-OCT-2000; 2000US-0240117P.

XX (FERR) FERRING BV.

XX Qi S, Akinsanya KO, Riviere PJ, Junien J;

XX WPI; 2002-444178/47.

XX P-PSDB; ABG61600.

XX New dipeptidyl peptidase IV-related proteins and nucleic acids encoding
PT the proteins, useful for treating e.g. fungal, bacterial, protozoan and
PT viral infections, cancers, allergies, neurological disorders, or pain.

XX Disclosure; Page 72-73; 113pp; English.

XX The present invention relates to the isolation of novel human serine
CC proteases referred to as dipeptidyl peptidase IV (DPPIV)-related proteins
CC (DPRP). The dipeptidyl peptidase IV-related proteins (DPRP) and nucleic
CC acids encoding them are useful for treating infections such as fungal,
CC bacterial, protozoan and viral infections, particularly infections caused
CC by human immunodeficiency virus (HIV-1 or HIV-2), pain, diabetes,
CC precocious puberty, infertility, obesity, anorexia, bulimia, Parkinson's
CC disease, acute heart failure, hypertension, hypertension, urinary
CC retention, osteoporosis, angina pectoris, stroke, ulcers, asthma,
CC allergies, cancers, migraine, vomiting, psychotic and neurological
CC disorders (e.g. anxiety, dementia, or schizophrenia), and dyskinasias.
CC These may also be used in discovering therapeutic agents for the
CC treatment of reproductive, inflammatory and metabolic disorders. ABK83322
CC -ABK83343 encode human DPRP proteins

XX Sequence 4676 BP; 1424 A; 859 C; 979 G; 1414 T; 0 U; 0 Other;

Query Match 89.9%; Score 2806.4; DB 6; Length 4676;
Best Local Similarity 95.1%; Pred. No. 0;
Matches 2968; Conservative 0; Mismatches 1; Indels 151; Gaps 1;

QY 1 AAGTGTAAAGCTCCGAGGCCGCTGCTACTGCGCCGCTGCTTCTTAGTCCG 60
Db 1 AAGTGTAAAGCTCCGAGGCCGCTGCTACTGCGCCGCTGCTTCTTAGTCCG 60

QY 61 CGTTCGCGCTGGTGTGTCACGGCGCGCGGAGGAGCACTGCAACACGAGGACCG 120

Db 61 CGTTCGCGCTGGTGTGTCACGGCGCGCGGAGGAGCACTGCAACACGAGGACCG 120
QY 121 GAGTGGAGCGCGCAGCATGAAGCGCGCAGCGCCGCTCCATAGCGCAGCTCGGACCG 180
Db 121 GAGTGGAGCGCGCAGCATGAAGCGCGCAGCGCCGCTCCATAGCGCAGCTCGGACCG 180
QY 181 TCGGGCGGGCGGGGAGGAAATGCAACATGCGCAGCAGCAATGGAACAGAACAG 240
Db 181 TCGGGCGGGCGGGGAGGAAATGCAACATGCGCAGCAGCAATGGAACAGAACAG 240
QY 241 CTGGGTGTTGAGATATTTGAAACTGCGGACTGTGAGGAGAAATATTGAATCA CAGGATCGG 300
Db 241 CTGGGTGTTGAGATATTTGAAACTGCGGACTGTGAGGAGAAATATTGAATCA CAGGATCGG 300
QY 301 CCTAAATGGAGCCCTTTTATGTTGAGCGGTATTCCTGGAGTCACTTAAAAAGCTCTT 360
Db 301 CCTAAATGGAGCCCTTTTATGTTGAGCGGTATTCCTGGAGTCACTTAAAAAGCTCTT 360
QY 361 GCGGATACAGAAATATCATGGCTACATGCTAGGCAAGGCAACCATGATTTTCACTGTTT 420
Db 361 GCGGATACAGAAATATCATGGCTACATGCTAGGCAAGGCAACCATGATTTTCACTGTTT 420
QY 421 GTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGAAATCTATTACCTTGGCATGCT 480
Db 421 GTGAAGAGGAATGATCCAGATGGACCTCATTCAGACAGAAATCTATTACCTTGGCATGCT 480
QY 481 GGTGAGACAGAGAAATACACTGTTTATTCGAAATTCGAAATCTCAATATAGACGA 540
Db 481 GGTGAGACAGAGAAATACACTGTTTATTCGAAATTCGAAATCTCAATATAGACGA 540
QY 541 GCAGTCTTAATGCTCTCTTTGGAAGCTCTTTTGGATCTTTTTCAGGCAACACTGGAAT 600
Db 541 GCAGTCTTAATGCTCTCTTTGGAAGCTCTTTTGGATCTTTTTCAGGCAACACTGGAAT 600
QY 601 GGAATGATTCGAGAGAGAACTATTAAGAGAGAAAGAAACGCAATTTGGAACAGTCCGA 660
Db 601 GGAATGATTCGAGAGAGAGAACTATTAAGAGAGAAAGAAACGCAATTTGGAACAGTCCGA 660
QY 661 ATTGCTTCTTAAGATTTATCAACAGAGAGTGAACATTTCTGTTTCAAGCGGTAGTGA 720
Db 661 ATTGCTTCTTAAGATTTATCAACAGAGAGTGAACATTTCTGTTTCAAGCGGTAGTGA 720
QY 721 ATTTATCAGTAAAGATGGAGGCGCACAGGATTTAGCAACACTTTTAAGGCCCAAT 780
Db 721 ATTTATCAGTAAAGATGGAGGCGCACAGGATTTAGCAACACTTTTAAGGCCCAAT 780
QY 781 CTAGTGGAACTAGTTGTCCTCAACATACGATGGATCCAAATTTATGCTGCTGATCCA 840
Db 781 CTAGTGGAACTAGTTGTCCTCAACATACGATGGATCCAAATTTATGCTGCTGATCCA 840
QY 841 GACTGATTCCTTTTATACATAGCAACGATATTTGGATATCTAACATCGTAACAGAGAA 900
Db 841 GACTGATTCCTTTTATACATAGCAACGATATTTGGATATCTAACATCGTAACAGAGAA 900
QY 901 GAAAGGAGACTCACTTATGTGCAATAGTAGTACCAACATGGAAGAGATGCCAGATCA 960
Db 901 GAAAGGAGACTCACTTATGTGCAATAGTAGTACCAACATGGAAGAGATGCCAGATCA 960
QY 961 GCTGGAGTCCCTACCTTTGTTCTCCAGAGAAATTTGATAGATATTTCTGGCTATTGGTGG 1020
Db 961 GCTGGAGTCCCTACCTTTGTTCTCCAGAGAAATTTGATAGATATTTCTGGCTATTGGTGG 1020
QY 1021 TGTCAAAAGCTGAAACCACTCCCGAGTGGTGGTAAATTTCTTGAATTTCTATATGAAGAA 1080
Db 1021 TGTCAAAAGCTGAAACCACTCCCGAGTGGTGGTAAATTTCTTGAATTTCTATATGAAGAA 1080
QY 1081 AATGATGAATCTGAGGTGGAATTTATCATGTTACATCCCTTATGTTGGAACAAGAGG 1140
Db 1081 AATGATGAATCTGAGGTGGAATTTATCATGTTACATCCCTTATGTTGGAACAAGAGG 1140
QY 1141 GCAGATTCATTCGGTTATTCCTAAACAGGTACAGCAAAATCTTAAAGTCACTTTTAAAGATG 1200
Db 1141 GCAGATTCATTCGGTTATTCCTAAACAGGTACAGCAAAATCTTAAAGTCACTTTTAAAGATG 1200

```
QY 1201 TCAGAAATAATGATTGATGCTGAAGGAGGATCATAGATGCTATAGATAAGGAATTAAT 1260
DB 1201 TCAGAAATAATGATTGATGCTGAAGGAGGATCATAGATGCTATAGATAAGGAATTAAT 1260
QY 1261 CAACCTTTTGAGATCTATTTTGAAGGAGTTGAATATATATGCGAGAGCTGGATGACTCT 1320
DB 1261 CAACCTTTTGAGATCTATTTTGAAGGAGTTGAATATATATGCGAGAGCTGGATGACTCT 1320
QY 1321 GAGGAAATAATGCTTGGTCCATCTACTAGATGCTCCAGACTCGCCCTACAGATAGTG 1380
DB 1321 GAGGAAATAATGCTTGGTCCATCTACTAGATGCTCCAGACTCGCCCTACAGATAGTG 1380
QY 1381 TTGATCTCACTCGAATATTTATCCCGTAGAAGATGATGTTATGGAAGGCGAGAGCTC 1440
DB 1381 TTGATCTCACTCGAATATTTATCCCGTAGAAGATGATGTTATGGAAGGCGAGAGCTC 1440
QY 1441 ATTGAGTCAGTGCCTGATTTCTGTGACGCCCACTAATTTATCTATGAAGAAACAAACAGATC 1500
DB 1441 ATTGAGTCAGTGCCTGATTTCTGTGACGCCCACTAATTTATCTATGAAGAAACAAACAGATC 1500
QY 1501 TGGATAAATATCCATGACATCTTTTCAATGTTTTTCCCAAGTCAACGAAGGAAATTTGAG 1560
DB 1501 TGGATAAATATCCATGACATCTTTTCAATGTTTTTCCCAAGTCAACGAAGGAAATTTGAG 1560
QY 1561 TTTATTTTGGCTCTGAAATGCAAAACAGGTTTCGGTCATTTATACAAAATTAACATCTATT 1620
DB 1561 TTTATTTTGGCTCTGAAATGCAAAACAGGTTTCGGTCATTTATACAAAATTAACATCTATT 1620
QY 1621 TTAAAGGAAGCAATATTAACGATCCAGTGGTGGCTGCCTCCCAAGTGAATTTCAAG 1680
DB 1621 TTAAAGGAAGCAATATTAACGATCCAGTGGTGGCTGCCTCCCAAGTGAATTTCAAG 1680
QY 1681 TGCTCTATCAAGAGGAGATAGCAATTAACAGTGGTGGATGGAAGTTCTTGGCGGGAT 1740
DB 1681 TGCTCTATCAAGAGGAGATAGCAATTAACAGTGGTGGATGGAAGTTCTTGGCGGGAT 1740
QY 1741 GGATCTAATATCCAAAGTTGATGAAGTCAAGAGCTGGTATATTTTGAAGGCCACCAAGAC 1800
DB 1741 GGATCTAATATCCAAAGTTGATGAAGTCAAGAGCTGGTATATTTTGAAGGCCACCAAGAC 1800
QY 1801 TCCCTTTTAGAGCATCACTGTAGTACGTAGTCAAGTAAATCTCGAGAGGTGACAAGG 1860
DB 1801 TCCCTTTTAGAGCATCACTGTAGTACGTAGTCAAGTAAATCTCGAGAGGTGACAAGG 1860
QY 1861 CTGACTGACCTGGCTACTCACATTTCTGTGCTGATCAGTCAGCACTGTGACTTTCTTTATA 1920
DB 1861 CTGACTGACCTGGCTACTCACATTTCTGTGCTGATCAGTCAGCACTGTGACTTTCTTTATA 1920
QY 1921 AGTAAGTATAGTAACCAAGAAATCCACACTGTGTGTCCTTTTACAAGCTATCAAGTCTT 1980
DB 1921 AGTAAGTATAGTAACCAAGAAATCCACACTGTGTGTCCTTTTACAAGCTATCAAGTCTT 1980
QY 1981 GAAGATGCCAACTTGCMAAACAAGGAATTTTGGGCCACCAATTTTGGATTCAGCAGGT 2040
DB 1981 GAAGATGCCAACTTGCMAAACAAGGAATTTTGGGCCACCAATTTTGGATTCAGCAGGT 2040
QY 2041 CCTCTTCTGACTATATCTCTCCAGAAATTTCTCTTTTGAAGATCACTACGATTTACA 2100
DB 2041 CCTCTTCTGACTATATCTCTCCAGAAATTTCTCTTTTGAAGATCACTACGATTTACA 2100
QY 2101 TTGATAGGATGCTCTACAAGCTCATGATCTACAGCTCGAAGAAATATCTCTACTGTG 2160
DB 2101 TTGATAGGATGCTCTACAAGCTCATGATCTACAGCTCGAAGAAATATCTCTACTGTG 2160
QY 2161 CTGTTTCATATATGTTGGTCTCTCAGGTGCTAGTGGTGAATATCGGTTTAAAGGAGTCAAG 2220
DB 2161 CTGTTTCATATATGTTGGTCTCTCAGGTGCTAGTGGTGAATATCGGTTTAAAGGAGTCAAG 2220
QY 2221 TATTTCCGCTTTGAATACCTTAGCCTCTCTAGGTTATGTGGTTATGTAGTAGACAAACAGG 2280
DB 2180 ----- 2179
```

```
QY 2281 GGATCCTGTCACGAGGGCTTAAATTTTGAAAGCGCCTTTAAATATATAAAATGGGTCAAATA 2340
DB 2180 -----GGGTCAAATA 2189
QY 2341 GAAATTTGACGATCAGGTGGAGGACTCCAATATCTAGCTTCTCGATATGATTTTCAATTGAC 2400
DB 2190 GAAATTTGACGATCAGGTGGAGGACTCCAATATCTAGCTTCTCGATATGATTTTCAATTGAC 2349
QY 2401 TTAGATCGTGTGGGATCAACGCGTGGTCTCTATGAGGATACCTCTCCCTGATGGCATTA 2460
DB 2250 TTAGATCGTGTGGGATCAACGCGTGGTCTCTATGAGGATACCTCTCCCTGATGGCATTA 2309
QY 2461 ATGAGAGTCCAGATATCTTTCAGGGTGTCTATTCCTGGGGCCCAAGTCACTCTGTGGATC 2520
DB 2310 ATGAGAGTCCAGATATCTTTCAGGGTGTCTATTCCTGGGGCCCAAGTCACTCTGTGGATC 2369
QY 2521 TTCTATGATACAGGATACACGGAACGTTATATGCGGTCAACCTGACCCAGAAATGAACAGGGC 2580
DB 2370 TTCTATGATACAGGATACACGGAACGTTATATGCGGTCAACCTGACCCAGAAATGAACAGGGC 2429
QY 2581 TATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGAACCAAATCGTTTA 2640
DB 2430 TATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGAACCAAATCGTTTA 2489
QY 2641 CTGCTCTTACATGGTTTCTCTGATGAGAAATGTCCTTTCGACATACCATATATTACTG 2700
DB 2490 CTGCTCTTACATGGTTTCTCTGATGAGAAATGTCCTTTCGACATACCATATATTACTG 2549
QY 2701 AGTTTTTTAGTGAGGGCTGGAAGCCCATATGATTTACAGATCTCTCTCAGGAGACAC 2760
DB 2550 AGTTTTTTAGTGAGGGCTGGAAGCCCATATGATTTACAGATCTATCTCTCAGGAGACAC 2609
QY 2761 AGCATAGAGTTCTTGAATCGGAGACATTAAGCACTGCATCTTTTGAACACTCTTCAA 2820
DB 2610 AGCATAGAGTTCTTGAATCGGAGACATTAAGCACTGCATCTTTTGAACACTCTTCAA 2669
QY 2821 GAAACCTTTGATACGATTTGCTGCTCTAAAAGTATATAATTTTGACCTGTGTGAAC 2880
DB 2670 GAAACCTTTGATACGATTTGCTGCTCTAAAAGTATATAATTTTGACCTGTGTGAAC 2729
QY 2881 TCTCTGGTATACACTGGCTATTTTAAACCAATGAGGAGGTTTAAATCAACAGAAAACACAGA 2940
DB 2730 TCTCTGGTATACACTGGCTATTTTAAACCAATGAGGAGGTTTAAATCAACAGAAAACACAGA 2789
QY 2941 ATTGATCATCAATTTTGATACCTGCGCATGTAAACATCTACTCTGAAAAATAAATGTGGTG 3000
DB 2790 ATTGATCATCAATTTTGATACCTGCGCATGTAAACATCTACTCTGAAAAATAAATGTGGTG 2849
QY 3001 CCATGACGGGTCTACGGTTTGTGTAGTAAATCTAAATACCTTAAACCCACATGCTCAAAA 3060
DB 2850 CCATGACGGGTCTACGGTTTGTGTAGTAAATCTAAATACCTTAAACCCACATGCTCAAAA 2909
QY 3061 TCAATATGATACATATTCCTGAGAGACCCAGCAATACCATTAAGATTAATAAAAAAAA 3120
DB 2910 TCAATATGATACATATTCCTGAGAGACCCAGCAATACCATTAAGATTAATAAAAAAAA 2969
```

RESULT 8

ACA92421 standard; DNA; 2952 BP.

XX ACA92421;

AC ACA92421;

DT 15-JUL-2003 (first entry)

XX DNA encoding human PMM-6.

XX Human; protein modification and maintenance molecule; PMM; cancer;

KW cell proliferation disorder; atherosclerosis; neurological disorder;

KW epilepsy; Huntington's disease; stroke; immune disorder; allergy;

KW inflammatory disorder; AIDS; developmental disorder; hypothyroidism;

KW Cushing's syndrome; gastrointestinal disorder; epithelial disorder;

KW infection; cytostatic; antiarteriosclerotic; anticonvulsant; nootropic;

neuroprotective; cerebroprotective; anti-HIV; antiallergic; vulnerary;
antiinflammatory; thyromimetic; gene; ds.

Homo sapiens.

WO2003031939-A2.

17-APR-2003.

11-OCT-2002; 2002WO-US032850.

12-OCT-2001; 2001US-0329689P.

25-OCT-2001; 2001US-0335703P.

09-NOV-2001; 2001US-0348887P.

28-NOV-2001; 2001US-0334145P.

06-DEC-2001; 2001US-0337451P.

14-DEC-2001; 2001US-0340584P.

(INCY-) INCYTE GENOMICS INC.

Ramkumar J, Gorvad AE, Baughn MR, Emerling BM, Yang J, Lee SY;

Tran UK, Becha SD, Duggan BM, Lee EA, Griffin JA, Li JX;

Sprague WW, Hatalia AJA, Chawla NK, Lehr-Mason PM, Kable AE, Yue H;

Marquis JP, Yao MG, Richardson TW, Tang TY, Jin P, Chien D;

Bhatia U, Burrill JD, Lee S, Blake JJ, Ho A, Zheng W;

WPI; 2003-430274/40.

P-PSDB; AB092026.

New human protein modification and maintenance molecules (PMM), useful
for diagnosing, treating and preventing diseases or conditions associated
with the aberrant PMM expression e.g. cancer, atherosclerosis, or
infections.

Claim 5; Page 285-286; 31pp; English.

The present invention relates to the isolation of human protein
modification and maintenance molecules (PMM), and the polynucleotide
sequences encoding them. A total of 40 PMM polypeptides (designated PMM
-1 to PMM-40) are disclosed. The sequences of the invention are useful
for diagnosing a condition or disease associated with the expression of
PMM in a subject, preparing a polyclonal or monoclonal antibody, and
generating an expression profile of a sample containing the
polynucleotides. The diseases or conditions associated with decreased
expression or overexpression of PMM are cell proliferation disorders
(e.g. cancer, atherosclerosis), neurological disorders (e.g. epilepsy,
Huntington's disease, stroke), immune/inflammatory disorders, (e.g. AIDS,
allergies), developmental disorders (e.g. hypothyroidism, Cushing's
syndrome), gastrointestinal or epithelial disorders, and infections. The
PMM polypeptides or their fragments are useful in screening compounds
for effectiveness as agonists or antagonists of the polypeptides, or in
altering the expression of the target polynucleotide and compounds that
specifically bind to, or modulate the activity of the polypeptide.

ACA92416-ACA92455 encode the human PMM polypeptides of the invention

Sequence 2952 BP; 886 A; 604 C; 664 G; 798 T; 0 U; 0 Other;
Query Match 89.3%; Score 2785.8; DB 10; Length 2952;
Best Local Similarity 95.0%; Pred. No. 0;
Matches 2950; Conservative 0; Mismatches 2; Indels 153; Gaps 1;

11 GCCTCGAGGCCAAGCGCGCTGCTACTGCGCGCGCTGCTTCTAGTGGCGTTCGCCGC 70
1 GCCTCGAGGCCAAGCGCGCTGCTACTGCGCGCGCTGCTTCTAGTGGCGTTCGCCGC 60

71 CTGGGTGTGTCACCGCGCGCGCGCGAGGAGCCACTGCAACACGAGACCGGAGTGGAGGC 130

61 CTGGGTGTGTCACCGCGCGCGCGCGAGGAGCCACTGCAACACGAGACCGGAGTGGAGGC 120

131 GCGCGACGATGAAGCGCGCGCGCGCGCTCCATAGCGCACGTGGGACGCTCGCGCGCGG 190

121 GCGCGACGATGAAGCGCGCGCGCGCGCTCCATAGCGCACGTGGGACGCTCGCGCGCGG 180

191 GCCGGGGGGAAGGAAATGCAACATGCGCAGCAATGGAACAGAAACAGCTGGGTGTTG 250
181 GCCGGGGGGAAGGAAATGCAACATGCGCAGCAATGGAACAGAAACAGCTGGGTGTTG 240
251 AGATATTTGAAACTGCGGACTGTGAGGAGAAATTTGAATCACAGGATCGGCCCTAAATTTG 310
241 AGATATTTGAAACTGCGGACTGTGAGGAGAAATTTGAATCACAGGATCGGCCCTAAATTTG 300
311 AGCCTTTTATGTTGAGCGGTATTTCTTGGAGTCACTTAAAGAGCTGCTTGGCGATACCA 370
301 AGCCTTTTATGTTGAGCGGTATTTCTTGGAGTCACTTAAAGAGCTGCTTGGCGATACCA 360
371 GAAATATACATGCTTACATGATGCTAGGCAACACCACTATTTTCATCTTTGTGAAGAGGA 430
361 GAAATATACATGCTTACATGATGCTAGGCAACACCACTATTTTCATCTTTGTGAAGAGGA 420
431 ATGATCCAGATGGAACCTCATTTACAGACAAATCTATTACCTTTGCCATGCTTGGTGAGAAC 490
421 ATGATCCAGATGGAACCTCATTTACAGACAAATCTATTACCTTTGCCATGCTTGGTGAGAAC 480
491 GAGAAATACACTGTTTATTTCTGAAATTCCTCAAACTATCATATAGAGAGCAGTCTTAA 550
481 GAGAAATACACTGTTTATTTCTGAAATTCCTCAAACTATCATATAGAGAGCAGTCTTAA 540
551 TGCTCTCTTGGAGGCTCTTTTGGATCTTTTTCAGGCAACACTGCACTATGGAATGTTT 610
541 TGCTCTCTTGGAGGCTCTTTTGGATCTTTTTCAGGCAACACTGCACTATGGAATGTTT 600
611 CTCGAGAGAGAACTTATTAGAGAGAGAAACGATTTGGAAACAGTCCGGAATTCCTTCTT 670
601 CTCGAGAGAGAACTTATTAGAGAGAGAAACGATTTGGAAACAGTCCGGAATTCCTTCTT 660
671 ACGATTATCACCAGAGAGAGTGGAACTTTCTTGGTTTCAAGCGGTAGTGGAAATTTATCAG 730
661 ACGATTATCACCAGAGAGAGTGGAACTTTCTTGGTTTCAAGCGGTAGTGGAAATTTATCAG 720
731 TAAAGATGAGGGGCCACAGGATTTAGCAACACCTTTTAAAGCCCAATCTAGTGGAAA 790
721 TAAAGATGAGGGGCCACAGGATTTAGCAACACCTTTTAAAGCCCAATCTAGTGGAAA 780
791 CTAGTTGTCCTCAACATACGATGGAATCCAAATTTATGCCCGCTGATCCAGACTGGATTG 850
781 CTAGTTGTCCTCAACATACGATGGAATCCAAATTTATGCCCGCTGATCCAGACTGGATTG 840
851 CTTTTATACATAGCAACGATATTTGGATATCTAACATCTTAACAGAGAGAGAGAGAC 910
841 CTTTTATACATAGCAACGATATTTGGATATCTAACATCTTAACAGAGAGAGAGAGAC 900
911 TCATTTATGTCACATAGAGCTAGCCCAACATGGAAGAGATGCCAGATCAGCTGGAGTCG 970
901 TCATTTATGTCACATAGAGCTAGCCCAACATGGAAGAGATGCCAGATCAGCTGGAGTCG 960
971 CTACCTTTGTTCTCCAGAGAGATTTGATAGATATTTCTGGCTATTGGTGGTGTCCAAAG 1030
961 CTACCTTTGTTCTCCAGAGAGATTTGATAGATATTTCTGGCTATTGGTGGTGTCCAAAG 1020
1031 CTGAAAACAATCTCCAGTGGTGTGAAATTTCTTAGAATTTCTATATGAAGAGAAATGATGAAT 1090
1021 CTGAAAACAATCTCCAGTGGTGTGAAATTTCTTAGAATTTCTATATGAAGAGAAATGATGAAT 1080
1091 CTGAGGTGGAATTTATTTATGTTTACATCCCTTATGTTGGAACAAGAGAGGGGAGATTCAT 1150
1081 CTGAGGTGGAATTTATTTATGTTTACATCCCTTATGTTGGAACAAGAGAGGGGAGATTCAT 1140
1151 TCGTGTATCTTAAACAAGGTACAGCAATCTTAAAGTCACTTTTAAGATGTTCAGAAATPAA 1210
1141 TCGTGTATCTTAAACAAGGTACAGCAATCTTAAAGTCACTTTTAAGATGTTCAGAAATPAA 1200
1211 TGATTGATGCTGAAGAGAGATCATAGATGTCTATAGATAAGGAACTAATTTCAACCTTTTG 1270
1201 TGATTGATGCTGAAGAGAGATCATAGATGTCTATAGATAAGGAACTAATTTCAACCTTTTG 1260
1271 AGATTCTTATTGAAGAGAGTTGAATATATTGCGCAGAGCTGGATGGAGCTCTCTCTGAGGAAAT 1330

XX (INCY-) INCYTE CORP.
XX Furness LM, Buchbinder JL;
XX WPI; 2004-068610/07.
XX Combination useful for preparing a composition for treating liver
XX disorders associated with steroid therapy, e.g., cirrhosis or hepatitis,
XX comprises cDNAs that are differentially expressed in response to steroid
XX treatment.
XX Claim 1; SEQ ID NO 1103; 141pp; English.
XX The invention relates to a combination comprising cDNAs that are
XX differentially expressed in response to steroid treatment. Also included
XX are the following: a high throughput method for using a cDNA to detect
XX differential expression of nucleic acids in a sample; and a high
XX throughput method of screening molecules or compounds to identify a
XX ligand that specifically binds a cDNA. The sample is from a subject with
XX Wilson disease and comparison of a standard defines a stage of that
XX disease. The high throughput method of screening molecules or compounds
XX to identify a ligand that specifically binds a cDNA comprises: combining
XX the combination with molecules or compounds under conditions to allow
XX specific binding; and detecting specific binding between each cDNA and at
XX least one molecule or compound. The molecules or compounds are regulatory
XX proteins. The combination is useful for preparing a composition for
XX treating liver disorders associated with steroid therapy, e.g., cirrhosis
XX or hepatitis. The present sequence represents a human cDNA which is
XX differentially expressed in steroid-induced C3A liver cells. Note: The
XX sequence data for this patent did not form part of the printed
XX specification but was obtained in electronic format directly from USPTO
XX at seqdata.uspto.gov/sequence.html.
SQ Sequence 2797 BP; 866 A; 551 C; 595 G; 785 T; 0 U; 0 Other;
Query Match 88.48; Score 2759.4; DB 12; Length 2797;
Best Local Similarity 99.8; Pred. No. 0;
Matches 2785; Conservative 0; Mismatches 1; Indels 5; Gaps 2;
QY 335 CCTGGAGTCAGCTTAAAGCTGCTTGGCCGATACAGAGAAATATCATGGCTACATGATGG 394
DB 1 CCTGGAGTCAGCTTAAAGCTGCTTGGCCGATACAGAGAAATATCATGGCTACATGATGG 60
QY 395 CTAAGGCACACATGATTTTCATGTTGTGAAGAGGAATGATCCAGATGGACCTCATTCAG 454
DB 61 CTAAGGCACACATGATTTTCATGTTGTGAAGAGGAATGATCCAGATGGACCTCATTCAG 120
QY 455 ACAGAACTCTATTACCTTGGCATGTCTGGTGAGAACAGAGAAATACATGTTTATTCTG 514
DB 121 ACAGAACTCTATTACCTTGGCATGTCTGGTGAGAACAGAGAAATACATGTTTATTCTG 180
QY 515 AAATTCGCAAACTATCAATAGACGACGTCTTAATGCTCTCTTGGAGCTCTTTTGG 574
DB 181 AAATTCGCAAACTATCAATAGACGACGTCTTAATGCTCTCTTGGAGCTCTTTTGG 240
QY 575 ATCTTTTCAGGCAACACTGACTATGGAATGATTTCTCGAGAGAGAACTATTAAAGAG 634
DB 241 ATCTTTTCAGGCAACACTGACTATGGAATGATTTCTCGAGAGAGAACTATTAAAGAG 300
QY 635 AAGAGAAACGATTTGGACACAGTCGGAATTCCTTTACGATTTATCACCAGAGAGTGGAA 694
DB 301 AAGAGAAACGATTTGGACACAGTCGGAATTCCTTTACGATTTATCACCAGAGAGTGGAA 360
QY 695 CATTTCTCTTTCAAGCCGCTAGTGGAAATTTATCACGTAAGAGATGGAGGCCCAAGGAT 754
DB 361 CATTTCTCTTTCAAGCCGCTAGTGGAAATTTATCACGTAAGAGATGGAGGCCCAAGGAT 420
QY 755 TTACGCAACACCTTTTAAGGCCCAATCTAGTGGAAACTAGTTGTCCCAACATACCGATGG 814
DB 421 TTACGCAACACCTTTTAAGGCCCAATCTAGTGGAAACTAGTTGTCCCAACATACCGATGG 480
QY 815 ATCCAAATATATGCCCCGCTGATCCAGACTGGATGCTTTTATACATAGCAACGATATTT 874

DB 481 ATCCAAATATATGCCCCGCTGATCCAGACTGGATGCTTTTATACATAGCAACGATATTT 540
QY 875 GGATATCTACATCGTAACACAGAGAGAAAGAGAGACTCATTATGTGACAAATGAGCTAG 934
DB 541 GGATATCTACATCGTAACACAGAGAGAAAGAGAGACTCATTATGTGACAAATGAGCTAG 600
QY 935 CCAACATGGAAGATGCGCAGATCAGCTGGAGTGGCTACCTTTGTTCTTCCAAAGAGAT 994
DB 601 CCAACATGGAAGATGCGCAGATCAGCTGGAGTGGCTACCTTTGTTCTTCCAAAGAGAT 660
QY 995 TTGATAGATATCTGGCTATTGGTGGTCCAAAAGCTGAAACAACTCCAGTGGTGGTA 1054
DB 661 TTGATAGATATCTGGCTATTGGTGGTCCAAAAGCTGAAACAACTCCAGTGGTGGTA 720
QY 1055 AAATCTTAGAATCTTATAGAAAGAAATGATGATCTGAGGTGGAATTTATCATGTTA 1114
DB 721 AAATCTTAGAATCTTATAGAAAGAAATGATGATCTGAGGTGGAATTTATCATGTTA 780
QY 1115 CATCCCTATGTTGGAACCAAGG-AGGCGAGATTCATTCCGTTATCTTAAACAGGTACA 1173
DB 781 CATCCCTATGTTGGAACCAAGGCGAGGCGAGATTCATTCCGTTATCTTAAACAGGTACA 840
QY 1174 GCAATCTTAAGTCACTTTTAAGATGTCAGAAATAATGATGATCTGAGAGAGGTGAA 1233
DB 841 GCAATCTTAAGTCACTTTTAAGATGTCAGAAATAATGATGATCTGAGAGAGGTGAA 900
QY 1234 ATAGATGTCATAGATAAGGAATCAATCAACCTTTTGAAGATTCATTGAGAGAGGTGAA 1293
DB 901 ATAGATGTCATAGATAAGGAATCAATCAACCTTTTGAAGATTCATTGAGAGAGGTGAA 960
QY 1294 TATATTGCCAGAGCTGGATGGACTCTCTGAGGGAATA----TGCTTGGTCCATCTACT 1349
DB 961 TATATTGCCAGAGCTGGATGGACTCTCTGAGGGAATAATGATGCTGCTGCTCCTACT 1020
QY 1350 AGATCGCTCCAGACTCGCCTACAGATGTTGATCTCACTGAAATTTATTTATCCAGT 1409
DB 1021 AGATCGCTCCAGACTCGCCTACAGATGTTGATCTCACTGAAATTTATTTATCCAGT 1080
QY 1410 AGAAGATGATGTTATGGAAGGAGAGACTCATTGAGTCAAGTCTGATCTCTGACGCC 1469
DB 1081 AGAAGATGATGTTATGGAAGGAGAGACTCATTGAGTCAAGTCTGATCTCTGACGCC 1140
QY 1470 ACTAATTTATCTTGAAGAAACACAGACATCTGGATAAATATCCATGATCTTTCACTG 1529
DB 1141 ACTAATTTATCTTGAAGAAACACAGACATCTGGATAAATATCCATGATCTTTCACTG 1200
QY 1530 TTTTCCCAAGTCAAGAGAGAAATTTGAGTTTATTTTGGCTCTGTAATGCAAAACAGG 1589
DB 1201 TTTTCCCAAGTCAAGAGAGAAATTTGAGTTTATTTTGGCTCTGTAATGCAAAACAGG 1260
QY 1590 TTTCCGTCATTTATACAAATTTACATCTATTTTAAAGGAAGCAATATAAAGCATCAG 1649
DB 1261 TTTCCGTCATTTATACAAATTTACATCTATTTTAAAGGAAGCAATATAAAGCATCAG 1320
QY 1650 TGGTGGGCTGCTCTCCAAAGTGAATTTCAAGTCTCTATCAAGAGAGAGATAGCAATAC 1709
DB 1321 TGGTGGGCTGCTCTCCAAAGTGAATTTCAAGTCTCTATCAAGAGAGAGATAGCAATAC 1380
QY 1710 CAGTGGTGAATGGGAAGTCTTGGCGGCGATGATCTAATATATCAAGTGTGATGAGTACG 1769
DB 1381 CAGTGGTGAATGGGAAGTCTTGGCGGCGATGATCTAATATATCAAGTGTGATGAGTACG 1440
QY 1770 AAGGCTGTATATTTTGAAGCACCAGAGACTCCCTTTTAGAGCATCACTGTACGTAGT 1829
DB 1441 AAGGCTGTATATTTTGAAGCACCAGAGACTCCCTTTTAGAGCATCACTGTACGTAGT 1500
QY 1830 CAGTTACCTAAATCTCGAGAGGTGACAGGCTGATGACCGTGGCTACTCTCACTTTCTG 1889
DB 1501 CAGTTACCTAAATCTCGAGAGGTGACAGGCTGATGACCGTGGCTACTCTCACTTTCTG 1560
QY 1890 CTGCAATCAGTCAAGCTGCTGACTCTTTTATAGTATAGTAAACCAAGAGATCCACA 1949

CC allergies), developmental disorders (e.g. hypothyroidism, Cushing's
 CC syndrome), gastrointestinal or epithelial disorders, and infections. The
 CC PMM polypeptides or their fragments are useful in screening compounds
 CC for effectiveness as agonists or antagonists of the polypeptides, or in
 CC altering the expression of the target polynucleotide and compounds that
 CC specifically bind to, or modulate the activity of the polypeptide.
 CC ACA92416-ACA92455 encode the human PMM polypeptides of the invention
 XX

SQ Sequence 2929 BP; 864 A; 606 C; 661 G; 798 T; 0 U; 0 Other;

Query Match 87.9%; Score 2742.4; DB 10; Length 2929;
 Best Local Similarity 94.4%; Pred. No. 0;
 Matches 2927; Conservative 0; Mismatches 1; Indels 174; Gaps 1;

QY 13 CTCGAGGCCAAGCGCGCTGCTACTGCGCGCGCTGCTTCTAGTGGCGGTTCCGCGCT 72
 DB |||||||
 QY 2 CTCGAGGCCAAGCGCGCTGCTACTGCGCGCGCTGCTTCTAGTGGCGGTTCCGCGCT 61
 DB |||||||
 QY 73 GGGTTGTCAACGCGCGCGCGCGAGGAAGCCACTGCAACACAGGACCGGAGTGGAGCGG 132
 DB |||||||
 QY 62 GGGTTGTCAACGCGCGCGCGCGAGGAAGCCACTGCAACACAGGACCGGAGTGGAGCGG 121
 DB |||||||
 QY 133 CGCAGCATGAAGCGCGCGCGCTCCATAGCGCAAGTGGAGCGGTCGGGCGGGGC 192
 DB |||||||
 QY 122 CGCAGCATGAAGCGCGCGCGCGCTCCATAGCGCAAGTGGAGCGGTCGGGCGGGC 181
 DB |||||||
 QY 193 CGGGGGAAGAAATGCAACATGGCAGCAGCAATGGAACAGAAACAGTGGGTGTGAG 252
 DB |||||||
 QY 182 CGGGGGAAGAAATGCAACATGGCAGCAGCAATGGAACAGAAACAGTGGGTGTGAG 241
 DB |||||||
 QY 253 ATATTTGAAATGCGGACTGTGAGGAAATATTTGAATACAGGATCGGCTTAAATGGAG 312
 DB |||||||
 QY 242 ATATTTGAAATGCGGACTGTGAGGAAATATTTGAATACAGGATCGGCTTAAATGGAG 301
 DB |||||||
 QY 313 CTTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGTCTGCTCCGATACCAGA 372
 DB |||||||
 QY 302 CTTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGTCTGCTCCGATACCAGA 361
 DB |||||||
 QY 373 AAATATCATGGCTACATGAGGCTAAGGACCAATGATTTTCTGTTGAGGAGGAT 432
 DB |||||||
 QY 362 AAATATCATGGCTACATGAGGCTAAGGACCAATGATTTTCTGTTGAGGAGGAT 421
 DB |||||||
 QY 433 GATCCAGATGGAACCTCATTCAGACAGAAATCTATTACCTGCCATGCTGGTGAGAACAGA 492
 DB |||||||
 QY 422 GATCCAGATGGAACCTCATTCAGACAGAAATCTATTACCTGCCATGCTGGTGAGAACAGA 481
 DB |||||||
 QY 493 GAAATACACTGTTTATTTCTGAAATCCCAAAATATCAATAGACAGCAGTCTTAATG 552
 DB |||||||
 QY 482 GAAATACACTGTTTATTTCTGAAATCCCAAAATATCAATAGACAGCAGTCTTAATG 541
 DB |||||||
 QY 553 CTCTCTTGGAAAGCCCTTTTGGATCTTTTTCAGGCAACACTGGACTATGGAATGATTTCT 612
 DB |||||||
 QY 542 CTCTCTTGGAAAGCCCTTTTGGATCTTTTTCAGGCAACACTGGACTATGGAATGATTTCT 573
 DB |||||||
 QY 613 CGAGAAGAAGAACTATTAAAGAAAGAAAGCAAGCAATGGAACACTGGAAATGCTTCTTAC 672
 DB |||||||
 QY 574 ----- 573
 QY 673 GATTATCACAAGGAAGTGGAAACATTTCTGTTTCAAGCGGTAGTGGAAATTTATCAGTGA 732
 DB |||||||
 QY 574 ----- 573
 QY 733 AAAGATGGAGGGCCCAAGGATTTACGCAACACTTTTAAGGCCCAATCTAGTGGAACT 792
 DB |||||||
 QY 574 -----GCAACAACTTTAAGGCCCAATCTAGTGGAACT 607
 DB |||||||
 QY 793 AGTTGTCCCAACATACGATGATGCCAAATTTATGCCCGCTGATCCAGACTGGATTGCT 852
 DB |||||||
 QY 608 AGTTGTCCCAACATACGATGATGCCAAATTTATGCCCTGCTGATCCAGACTGGATTGCT 667
 DB |||||||
 QY 853 TTTATACATAGCAAGATATTTGGATATCTAACTATGTAACAGTAAACAGAGAAAGAGACTC 912
 DB |||||||
 QY 668 TTTATACATAGCAAGATATTTGGATATCTAACTATGTAACAGTAAACAGAGAAAGAGACTC 727
 DB |||||||

QY 913 ACTTATGTGCACATGAGCTAGCCAAACATGGAAGAAGATCCAGATCAGCTGAGTCGCT 972
 DB |||||||
 QY 728 ACTTATGTGCACATGAGCTAGCCAAACATGGAAGAAGATCCAGATCAGCTGAGTCGCT 787
 DB |||||||
 QY 973 ACCTTTGTTCTCCCAAGAGAATTTTATAGATATTTCTGGCTATTTGGTGGTCTCCAAAGCT 1032
 DB |||||||
 QY 788 ACCTTTGTTCTCCCAAGAGAATTTTATAGATATTTCTGGCTATTTGGTGGTCTCCAAAGCT 847
 DB |||||||
 QY 1033 GAAACAACTCCCAAGTGGTGGTAAATTTCTTAGAATTTCTATATGAAGAAATATGATATCT 1092
 DB |||||||
 QY 848 GAAACAACTCCCAAGTGGTGGTAAATTTCTTAGAATTTCTATATGAAGAAATATGATATCT 907
 DB |||||||
 QY 1093 GAGGTGGAATATTTATCTGTTACATCCCCTATGTTGGAACAAGGAGGCGAGATTCTTTC 1152
 DB |||||||
 QY 908 GAGGTGGAATATTTATCTGTTACATCCCCTATGTTGGAACAAGGAGGCGAGATTCTTTC 967
 DB |||||||
 QY 1153 CGTTATCTTAAACAGGTACAGCAATCTTAAAGTCACTTTTAAAGATGTCAGAAATAATG 1212
 DB |||||||
 QY 968 CGTTATCTTAAACAGGTACAGCAATCTTAAAGTCACTTTTAAAGATGTCAGAAATAATG 1027
 DB |||||||
 QY 1213 ATTGATGCTGAAGGAAGGATCATAGATGTCATAGATAAGGAATTAATTTCAACCTTTTGG 1272
 DB |||||||
 QY 1028 ATTGATGCTGAAGGAAGGATCATAGATGTCATAGATAAGGAATTAATTTCAACCTTTTGG 1087
 DB |||||||
 QY 1273 ATTCTATTTGAAGGAGTTGAATATATTTGCGAGAGCTGGATGGAATCTCTGAGGGAATAT 1332
 DB |||||||
 QY 1088 ATTCTATTTGAAGGAGTTGAATATATTTGCGAGAGCTGGATGGAATCTCTGAGGGAATAT 1147
 DB |||||||
 QY 1333 GCTTGTGTCATCTCTACTAGATGCTCCAGAGCTGCTTACAGATAGTGTGATCTCACT 1392
 DB |||||||
 QY 1148 GCTTGTGTCATCTCTACTAGATGCTCCAGAGCTGCTTACAGATAGTGTGATCTCACT 1207
 DB |||||||
 QY 1393 GAATTTATTTATCCAGTGAAGATGATGTTATGGAAGGAGAGACTCAITTAGTCACTG 1452
 DB |||||||
 QY 1208 GAATTTATTTATCCAGTGAAGATGATGTTATGGAAGGAGAGACTCAITTAGTCACTG 1267
 DB |||||||
 QY 1453 CTTGATTTCTGTGACGCCACTAATTTATCTATGAAGAAACACAGACATCTGGATAAATATC 1512
 DB |||||||
 QY 1268 CTTGATTTCTGTGACGCCACTAATTTATCTATGAAGAAACACAGACATCTGGATAAATATC 1327
 DB |||||||
 QY 1513 CATGACATCTTTCTGTTTCCCAAGTCCAGAGGAAATTTGATTTATTTTGGC 1572
 DB |||||||
 QY 1328 CATGACATCTTTCTGTTTCCCAAGTCCAGAGGAAATTTGATTTATTTTGGC 1387
 DB |||||||
 QY 1573 TCTGATGCAAAACAGGTTTCGTCATTTATACAAAATTTACATCTTATTTAAAGGAAGC 1632
 DB |||||||
 QY 1388 TCTGATGCAAAACAGGTTTCGTCATTTATACAAAATTTACATCTTATTTAAAGGAAGC 1447
 DB |||||||
 QY 1633 AAATATAAACGATCCAGTGGTGGCTGCTCCCAAGTGAATTTCAAGTGTCTTATCAAA 1692
 DB |||||||
 QY 1448 AAATATAAACGATCCAGTGGTGGCTGCTCCCAAGTGAATTTCAAGTGTCTTATCAAA 1507
 DB |||||||
 QY 1693 GAGGAGATAGCAATTAACAGTGGTGAATGGGAGTTCTTGGCGGCGATGGATCTAATATC 1752
 DB |||||||
 QY 1508 GAGGAGATAGCAATTAACAGTGGTGAATGGGAGTTCTTGGCGGCGATGGATCTAATATC 1567
 DB |||||||
 QY 1753 CAAGTTGATGAGTCAAGAGGCTGGTATATTTTGAAGGCAACCAAGACTCCCTTTTAG 1812
 DB |||||||
 QY 1568 CAAGTTGATGAGTCAAGAGGCTGGTATATTTTGAAGGCAACCAAGACTCCCTTTTAG 1627
 DB |||||||
 QY 1813 CATCACTGTACAGTCAAGTGAATCTGTAAGTGAAGGCTGCAAGAGCTGACTGACCGT 1872
 DB |||||||
 QY 1628 CATCACTGTACAGTCAAGTGAATCTGTAAGTGAAGGCTGCAAGAGCTGACTGACCGT 1687
 DB |||||||
 QY 1873 GGCTACTCAATTTCTGCTGCATCAGTCAAGCTGCACTCTTCTTTAATGAATGATATGAT 1932
 DB |||||||
 QY 1688 GGCTACTCAATTTCTGCTGCATCAGTCAAGCTGCACTCTTCTTTAATGAATGATATGAT 1747
 DB |||||||
 QY 1933 AACAGAGAATCCACACTGTGTGCTTTTCAAGCTATCAAGTCTTGAAGATGACCCA 1992
 DB |||||||
 QY 1748 AACAGAGAATCCACACTGTGTGCTTTTCAAGCTATCAAGTCTTGAAGATGACCCA 1807
 DB |||||||

Db	241	 GACAGAATCTTATACCTTGCCATGCTGGTGAGAACAGAGAAAATACACTGTTTTATTCT	300
Qy	514	GAATTTCCAAAACATCAATAGACAGCAGCTTAATGCTCTCTTGAGACCTCTTTTG	573
Db	301	GAATTTCCAAAACATCAATAGACAGCAGCTTAATGCTCTCTTGAGACCTCTTTTG	360
Qy	574	GATCTTTTTCAGGCAACACTGGACTATGGAATGTATTTCTCGAGAAGAAGAACTATTAA	633
Db	361	GATCTTTTTCAGGCAACACTGGACTATGGAATGTATTTCTCGAGAAGAAGAACTATTAA	420
Qy	634	GAAGAAGACGATTTGGAAACAGTCGGAATTTGCTTTACGATATCAACCAAGGAGTGA	693
Db	421	GAAGAAGACGATTTGGAAACAGTCGGAATTTGCTTTACGATATCAACCAAGGAGTGA	480
Qy	694	ACATTTCTGTTTCAAGCCGGTAGTGGAAATTTATCACGTAAAAGATGGAGGCCACAAGGA	753
Db	481	ACATTTCTGTTTCAAGCCGGTAGTGGAAATTTATCACGTAAAAGATGGAGGCCACAAGGA	540
Qy	754	TTTACGCAACAACTTTAAGGCCCAACTCTAGTGGAAACTAGTTGTGCCAACAATACGGATG	813
Db	541	TTTACGCAACAACTTTAAGGCCCAACTCTAGTGGAAACTAGTTGTGCCAACAATACGGATG	600
Qy	814	GATCCAAAATTTATGCCCGCTGATCCAGACTGGATTTGCTTTTATACATAGCAACGATATT	873
Db	601	GATCCAAAATTTATGCCCGCTGATCCAGACTGGATTTGCTTTTATACATAGCAACGATATT	660
Qy	874	TGGATATCTAAACATCGTAACAGAGAAAGAGAGACTCACCTTATGTGCACAATGAGCTA	933
Db	661	TGGATATCTAAACATCGTAACAGAGAAAGAGAGACTCACCTTATGTGCACAATGAGCTA	720
Qy	934	GCCAAACATGGAAGAAGATGCCAGATCAGCTGGAGTTCGCTACCTTTGTTCTCCAGAAGAA	993
Db	721	GCCAAACATGGAAGAAGATGCCAGATCAGCTGGAGTTCGCTACCTTTGTTCTCCAGAAGAA	780
Qy	994	TTTGATAGATATTCTGGCTATTGTTGGTGTGTCAAAAGCTGAAAACAATCCCAAGTGGT	1053
Db	781	TTTGATAGATATTCTGGCTATTGTTGGTGTGTCAAAAGCTGAAAACAATCCCAAGTGGT	840
Qy	1054	AAAATTTCTAGATTTCTATGAAGAAATGATGATCTGAGGTGGAAATTTATTCATGTT	1113
Db	841	AAAATTTCTAGATTTCTATGAAGAAATGATGATCTGAGGTGGAAATTTATTCATGTT	900
Qy	1114	ACATCCCTTATGTTGAAAACAAGGAGGCGAGATTCATTTCCGTTATCTCTAAAACAAGGTACA	1173
Db	901	ACATCCCTTATGTTGAAAACAAGGAGGCGAGATTCATTTCCGTTATCTCTAAAACAAGGTACA	960
Qy	1174	GCAATCTCTAAAGTCACTTTTAAAGATGTCAGAAATTAATGATGTCTGAAAGGAAGATC	1233
Db	961	GCAATCTCTAAAGTCACTTTTAAAGATGTCAGAAATTAATGATGTCTGAAAGGAAGATC	1020
Qy	1234	ATAGATGTCATAGATAAGAACTAATTCACCTTTTGAGATCTTATTTGAAGGAGTTGAA	1293
Db	1021	ATAGATGTCATAGATAAGAACTAATTCACCTTTTGAGATCTTATTTGAAGGAGTTGAA	1080
Qy	1294	TATATTGCCAGAGCTGGATGGACTCTCTGAGGAGAAATATGCTTGGTCCATCTCTACTAGAT	1353
Db	1081	TATATTGCCAGAGCTGGATGGACTCTCTGAGGAGAAATATGCTTGGTCCATCTCTACTAGAT	1140
Qy	1354	CGCTCCAGAGCTCGCCTACAGATAGTGTGATCTCACCTGAAATTTATTTATCCCAAGTAGAA	1413
Db	1141	CGCTCCAGAGCTCGCCTACAGATAGTGTGATCTCACCTGAAATTTATTTATCCCAAGTAGAA	1200
Qy	1414	GATGATGTTATGAAAGGAGAGACTCATTTAGTGCAGTGCCTGATTTCTGTGAGCCCACTA	1473
Db	1201	GATGATGTTATGAAAGGAGAGACTCATTTAGTGCAGTGCCTGATTTCTGTGAGCCCACTA	1260
Qy	1474	ATTATCTATGAAACACACACATCTGGATAAATATCCATGACATCTTTTCATGTTTTT	1533
Db	1261	ATTATCTATGAAACACACACATCTGGATAAATATCCATGACATCTTTTCATGTTTTT	1320
Qy	1534	CCCCAAGTCAGGAAGGAAATTTGATGTTTTTTTTTTTGGCTCTGAAATGCAAAACAGGTTTC	1593

QY 2674 CATTTTGCACATACAGTATATTAAGTATTTTATAGTGGGCTGGAAAGCCATATGAT 2733
Db 2461 CATTTTGCACATACAGTATATTAAGTATTTTATAGTGGGCTGGAAAGCCATATGAT 2520
QY 2734 TTACAGATCTATCTCTAGGAGAGACACAGCATAAGAGTTCTTGAATCGGAGAACATTTAT 2793
Db 2521 TTACAGATCTATCTCTAGGAGAGACACAGCATAAGAGTTCTTGAATCGGAGAACATTTAT 2580
QY 2794 GAACTGCATCTTTTGGCACTACCTTCAAGAAACCTTTGGATCACGATTATGCTGCTCTAAAA 2853
Db 2581 GAACTGCATCTTTTGGCACTACCTTCAAGAAACCTTTGGATCACGATTATGCTGCTCTAAAA 2640
QY 2854 GTGATATAA 2862
Db 2641 GTGATATAA 2649

RESULT 12
ABK83322 standard; cDNA; 2671 BP.
XX
AC ABK83322;
XX
DT 12-AUG-2002 (first entry)
XX
DE cDNA encoding human DPPIV related serine protease DPPR-1.
XX
KW Human; serine protease; dipeptidyl peptidase IV-related protein; DPPR;
KW DPPIV; infection; human immunodeficiency virus; HIV-1; HIV-2; pain;
KW diabetes; infertility; obesity; anorexia; Parkinson's disease; stroke;
KW heart failure; hypertension; urinary retention; osteoporosis; cancer;
KW ulcer; allergy; cancer; psychotic disorder; neurological disorder;
KW dyskinesia; reproductive disorder; inflammatory disorder;
KW metabolic disorder; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200231134-A2.
XX
PD 18-APR-2002.
XX
PF 12-OCT-2001; 2001WO-US031874.
XX
PR 12-OCT-2000; 2000US-0240117P.
XX
PA (FERR) FERRING BV.
XX
PI Qi S, Akinsanya KO, Riviere PJ, Junien J;
XX
XX WPI; 2002-444178/47.
XX P-PSDB; ABG61591.
XX
XX New dipeptidyl peptidase IV-related proteins and nucleic acids encoding
PT the proteins, useful for treating e.g. fungal, bacterial, protozoan and
PT viral infections, cancers, allergies, neurological disorders, or pain.
XX
XX Claim 1; Page 53-54; 113pp; English.
XX
XX The present invention relates to the isolation of novel human serine
CC proteases referred to as dipeptidyl peptidase IV (DPPIV)-related proteins
CC (DPPR). The dipeptidyl peptidase IV-related proteins (DPPR) and nucleic
CC acids encoding them are useful for treating infections such as fungal,
CC bacterial, protozoan and viral infections, particularly infections caused
CC by human immunodeficiency virus (HIV-1 or HIV-2), pain, diabetes,
CC precocious puberty, infertility, obesity, anorexia, bulimia, Parkinson's
CC disease, acute heart failure, hypotension, hypertension, urinary
CC retention, osteoporosis, angina pectoris, stroke, ulcers, asthma,
CC allergies, cancers, migraine, vomiting, psychotic and neurological
CC disorders (e.g. anxiety, dementia, or schizophrenia), and dyskinesias.
CC These may also be used in discovering therapeutic agents for the
CC treatment of reproductive, inflammatory and metabolic disorders. ABK83322
CC -ABK83343 encode human DPPR proteins

XX SQ Sequence 2671 BP; 805 A; 524 C; 594 G; 748 T; 0 U; 0 Other;
Query Match 84.8%; Score 2644.4; DB 6; Length 2671;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2648; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 209 GCAACATCGCAGCAGCAATGGAACAGACAGCTGGGTGTGAGATATTTGAACTGCCG 268
Db 3 GTACCATCGCAGCAGCAATGGAACAGACAGCTGGGTGTGAGATATTTGAACTGCCG 62
QY 269 ACTGTGAGGAGAATATTGAATCACAGGATCGGCTAAATTTGGAGCCTTTTATGTTGAGC 328
Db 63 ACTGTGAGGAGAATATTGAATCACAGGATCGGCTAAATTTGGAGCCTTTTATGTTGAGC 122
QY 329 GGTATTTCTGAGTCAAGCTTAAAAAGCTTGGCCGATACAGAAATATCATGGCTACA 388
Db 123 GGTATTTCTGAGTCAAGCTTAAAAAGCTTGGCCGATACAGAAATATCATGGCTACA 182
QY 389 TGATGGCTAAGGCCACCATGATTTTCAATTTGTGAAGAGGAATGATCCAGATGGACCTC 448
Db 183 TGATGGCTAAGGCCACCATGATTTTCAATTTGTGAAGAGGAATGATCCAGATGGACCTC 242
QY 449 ATTTCAGACAGATCTTATTTACCTTGCCTATGCTGTGAGAGACAGAGAAATATCATCTGTTT 508
Db 243 ATTTCAGACAGATCTTATTTACCTTGCCTATGCTGTGAGAGACAGAGAAATATCATCTGTTT 302
QY 509 ATTCTGAATTTCCCAACTATCAATAGACGACAGCTTAAATGCTCTCTTGGAGCCCTC 568
Db 303 ATTCTGAATTTCCCAACTATCAATAGACGACAGCTTAAATGCTCTCTTGGAGCCCTC 362
QY 569 TTTTGGATCTTTTTCAGGCAACACCTGACATATGGAATGATTTCTCGAGAGAGAACTAT 628
Db 363 TTTTGGATCTTTTTCAGGCAACACCTGACATATGGAATGATTTCTCGAGAGAGAACTAT 422
QY 629 TAAGAGAAAGAAACGCAATTTGGAACAGTCGGAATTTGCTTACGATATATCACAGAGAA 688
Db 423 TAAGAGAAAGAAACGCAATTTGGAACAGTCGGAATTTGCTTACGATATATCACAGAGAA 482
QY 689 GTGGACATTTCTGTTTCAAGCCGCTAGTGGAAATTTATCACTGTAAGAGTGGAGGCGCAC 748
Db 483 GTGGACATTTCTGTTTCAAGCCGCTAGTGGAAATTTATCACTGTAAGAGTGGAGGCGCAC 542
QY 749 AAGGATTTACGCACACCACTTTAAGGCGCCCAATCTAGTGGAAACCTAGTTGTCACCAATAC 808
Db 543 AAGGATTTACGCACACCACTTTAAGGCGCCCAATCTAGTGGAAACCTAGTTGTCACCAATAC 602
QY 809 GGATGGATCCAAAAATTTATGCCCGCTGATCCAGACTGGATTTGCTTTTATACATAGCAACG 868
Db 603 GGATGGATCCAAAAATTTATGCCCGCTGATCCAGACTGGATTTGCTTTTATACATAGCAACG 662
QY 869 ATATTGGATATCTAACATCTGTAACAGAGAAAGAGAGACTCACTTATGTGCACAATG 928
Db 663 ATATTGGATATCTAACATCTGTAACAGAGAGAAAGAGAGACTCACTTATGTGCACAATG 722
QY 929 AGCTAGCCAAATGGAAGAGATGCCAGATCAGCTGGAGTGGCTTGTGTTCTTCCCAAG 988
Db 723 AGCTAGCCAAATGGAAGAGATGCCAGATCAGCTGGAGTGGCTTGTGTTCTTCCCAAG 782
QY 989 AGAATTTGATAGATATTTGCTTATTTGCTGTCGCAAGAGTGAACACTCCCAATG 1048
Db 783 AGAATTTGATAGATATTTGCTTATTTGCTGTCGCAAGAGTGAACACTCCCAATG 842
QY 1049 GTGGTAAAAATCTTGAATTTTATGAAAGAAATGATGATCTGAGTGGAGAAATATTC 1108
Db 843 GTGGTAAAAATCTTGAATTTTATGAAAGAAATGATGATCTGAGTGGAGAAATATTC 902
QY 1109 ATGTTACATCCCTATGTTGGAACCAAGAGGCGCAGATTCATTCGTTTATCTTAAACAG 1168
Db 903 ATGTTACATCCCTATGTTGGAACCAAGAGGCGCAGATTCATTCGTTTATCTTAAACAG 962
QY 1169 GTACAGCAATCTTAAAGTCACTTTTAAAGATGTGCAAAATATGATTGATGCTGAGAGAA 1228

963	GTACAGCAAAATCCTTAAGTCACTTTTAAGATGTGAGAAATATGATTGATGCTGAGGAA	1026		
Qy	GGATCATAGATGTCATAGATAAGGAACATAATCAACCTTTTGAGATTCATATTTCAAGGAG	1288		
	GGATCATAGATGTCATAGATAAGGAACATAATCAACCTTTTGAGATTCATATTTGAAGGAG		1082	
	TTGAATATATTGGCAGAGCTGGATGGACTCCTGAGGGGAAAATATGCTTTGGTCCATCCTAC			1348
	TTGAATATATTGGCCAGAGCTGGATGGACTCCTGAGGGGAAAATATGCTTTGGTCCATCCTAC			
TAGATCGCTCCAGACTCGCCTCAGATAGTGTGATCTCACCTCAATATTATTATCCAG	1408			
TAGATCGCTCCAGACTCGCCTCAGATAGTGTGATCTCACCTCAATATTATTATCCAG		1202		
TAGAAGATGATGTTATTGGAAGAGGAGAGACTCATTTGAGTCAGTGCCTGATTTCTGTGACGC			1468	
TAGAAGATGATGTTATTGGAAGAGGAGAGACTCATTTGAGTCAGTGCCTGATTTCTGTGACGC				1262
CACTAATTATCTATGAAGAAACAAAGACATCTGGATTAATATCAATGACATCTTTTCATG	1528			
CACTAATTATCTATGAAGAAACAAAGACATCTGGATTAATATCAATGACATCTTTTCATG		1322		
TTTTTTTCCCAAGTCAAGAGAGAAATTCAGTTTATTTTTCGCTCTCAATGCAAAAACAG			1588	
TTTTTTTCCCAAGTCAAGAGAGAAATTCAGTTTATTTTTCGCTCTCAATGCAAAAACAG				1382
TTTTTTTCCCAAGTCAAGAGAGAAATTCAGTTTATTTTTCGCTCTCAATGCAAAAACAG	1648			
GTTTTCCGTCATTTATACAAAATTACATCTATTTTAAAGGAAAGCAAAATATAAAGCATCCA		1442		
GTTTTCCGTCATTTATACAAAATTACATCTATTTTAAAGGAAAGCAAAATATAAAGCATCCA			1708	
GTGTTGGGCTGCCTGCTCCAAGTGAATTTCAAGTGTCTTATCAAGAGGAGATAGCAATTA				1502
GTGTTGGGCTGCCTGCTCCAAGTGAATTTCAAGTGTCTTATCAAGAGGAGATAGCAATTA	1768			
CCAGTGGTGAATGGGAAGTCTTCGCGCGCATCGATCAATATCAAGTCTGATGAGTCA		1562		
CCAGTGGTGAATGGGAAGTCTTCGCGCGCATCGATCAATATCAAGTCTGATGAGTCA			1828	
GAAAGTGGTATATTTTGAAGGCACAAAGACTCCCTTTTAGAGCATCACCTGTACGTAG				1622
GAAAGTGGTATATTTTGAAGGCACAAAGACTCCCTTTTAGAGCATCACCTGTACGTAG	1888			
TCAGTTACGTAAATCCTCGAGAGGTGACAAGGCTGACTGACCGTGGTACTCACATTCCT		1948		
TCAGTTACGTAAATCCTCGAGAGGTGACAAGGCTGACTGACCGTGGTACTCACATTCCT			1742	
GCTGCATCAGTCAGACATCTGACATCTTTTATAGTAAGTATAGTATACCAAGAGATCCAC				2008
GCTGCATCAGTCAGACATCTGACATCTTTTATAGTAAGTATAGTATACCAAGAGATCCAC	1802			
ACTGTGTGCCCTTTTACAAGTCTCAAGTCTCGAAGTACCCAACTTTGCAAAAACAAAGG		2068		
ACTGTGTGCCCTTTTACAAGTCTCAAGTCTCGAAGTACCCAACTTTGCAAAAACAAAGG			2128	
AAATTTTGGGCCACCAATTTTGGATTCAGCAGGTCCTCTCTGACTATATCTCTCCAGAAA				1862
AAATTTTGGGCCACCAATTTTGGATTCAGCAGGTCCTCTCTGACTATATCTCTCCAGAAA	2188			
TTTTTCTCTTTGAAAGTACTACTGGATTTACATTTGATATGGATGTCTTACAGGCTCATG		1922		
TTTTTCTCTTTTGAAGTACTACTGGATTTACATTTGATATGGATGTCTTACAGGCTCATG			2192	
ATCTACAGCCTCGGAAGAAATATCTTACTGTGCTTTTATATATATGTTGGTCTCAGGTGC				1982
ATCTACAGCCTCGGAAGAAATATCTTACTGTGCTTTTATATATATGTTGGTCTCAGGTGC	2248			
AGTTGGTGAATTAATCCGTTTAAAGAGGTCAAGTATTTTCGCTTTGAATACCCCTAGCCTCTC		2042		
AGTTGGTGAATTAATCCGTTTAAAGAGGTCAAGTATTTTCGCTTTGAATACCCCTAGCCTCTC			2308	
TAGGTTATGTGTTGATGATAGACAACAGGGGATCCTGTGACCGAGGGCTTAAATTTG				2102
TAGGTTATGTGTTGATGATAGACAACAGGGGATCCTGTGACCGAGGGCTTAAATTTG	2168			

QY	2309	AAGGGCGCCCTTTAAATATAAATATGAGTGGGTCAAAATAGAAAATTGCACGATCAGGTGGAAGACTCC	2362
DB	2103	AAGGGCGCCCTTTAAATATAAATATGAGTGGGTCAAAATAGAAAATTGCACGATCAGGTGGAAGACTCC	2162
QY	2369	AATACTAGCTTCCTCGATATGATGTTTCATTGACCTTAGCATCGTGGGCCATCCACGGCTGGT	2428
DB	2163	ATAATCTAGCTTCCTCGATATGATGTTTCATTGACCTTAGCATCGTGGGCCATCCACGGCTGGT	2222
QY	2429	CCTATGGAGGATACCTCTCCCTCGATGGCATTAAATGCAGAGGTGATGATATCTTCAGGGTTG	2488
DB	2223	CCTATGGAGGATACCTCTCCCTCGATGGCATTAAATGCAGAGGTGATGATATCTTCAGGGTTG	2282
QY	2489	CTATTGCTGGGGCCCCAGTGCACTCTGTGGATCTTCTATGATCAGGATACACGGAACGTT	2548
DB	2283	CTATTGCTGGGGCCCCAGTGCACTCTGTGGATCTTCTATGATCAGGATACACGGAACGTT	2342
QY	2549	ATATGGGTCAACCTTGACCAGAGAAGACAGAGGCTATTACTTAGGATCTGTGGCCATGCAG	2608
DB	2343	ATATGGGTCAACCTTGACCAGAGAAGACAGAGGCTATTACTTAGGATCTGTGGCCATGCAG	2402
QY	2609	CAGAAAAAGTTCCCCTCTGGAACCAAATCGTTTTACTGCTCTTTACATGGTTTTCTCGATGAGA	2668
DB	2403	CAGAAAAAGTTCCCCTCTGGAACCAAATCGTTTTACTGCTCTTTACATGGTTTTCTCGATGAGA	2462
QY	2669	ATGTCATNTTGGCACATACACGATATATTAATCTAGATTTTTTTAGTAGGGCTGGAAGCCAT	2728
DB	2463	ATGTCATNTTGGCACATACACGATATATTAATCTAGATTTTTTTAGTAGGGCTGGAAGCCAT	2522
QY	2729	ATGATNTTACAGATCTATCTCTCGAGAGAGACACAGCATTAAGAGTTCCTGAATCGGAGAAC	2788
DB	2523	ATGATNTTACAGATCTATCTCTCGAGAGAGACACAGCATTAAGAGTTCCTGAATCGGAGAAC	2582
QY	2789	ATTATGAACCTGCATCTTTTGCACCTACCTTCAAGAAAACCTTGGATCAGGATATTCGCTGTC	2848
DB	2583	ATTATGAACCTGCATCTTTTGCACCTACCTTCAAGAAAACCTTGGATCAGGATATTCGCTGTC	2642
QY	2849	TAAAGTGATATTA 2862	
DB	2643	TAAAGTGATATGA 2656	
 RESULT 13 ABX12255			
ID	ABX12255 standard; cDNA; 2649 BP.		
XX	AC	ABX12255;	
XX	DT	19-MAY-2003 (first entry)	
XX	DE	cDNA encoding human serine protease HIPHUM46.	
XX	KW	Human; ss; gene; HIPHUM46; serine protease; gene therapy; osteoarthritis;	
KW	KW	serine protease activity modulation; dipeptidyl peptidase activity;	
KW	KW	musculoskeletal disease; Hepatitis B virus infection; myotonic dystrophy;	
KW	KW	Alzheimer's disease; paraspranuclear palsy; Huntington's disease;	
KW	KW	amyotrophic lateral sclerosis; malabsorption syndrome; lung disease;	
KW	KW	irritable bowel syndrome; type I diabetes; faecal incontinence; anaemia;	
KW	KW	haemorrhoid; proctitis; rectal polyp; small bowel tumour; dyslexia;	
KW	KW	colorectal tumour; ceroid lipofuscinosis; allergic encephalomyelitis;	
KW	KW	multiple sclerosis; chromosome 15q21-q22.	
XX	OS	Homo sapiens.	
XX	FH	Key Location/Qualifiers	
FT	CDS	1..2649	
FT	FT	product= "HIPHUM46"	
FT	FT	/note= "Serine protease"	
XX	GB	274869-A.	
PW	XX		
PD	30-OCT-2002.		

Tue Apr 18 08:18:37 2006

XX	22-JAN-2002; 2002GB-00001404.	QY	634	GAAGAAACGCGATTGGAAACAGTCGGAATTCCTTCTTACGATTATCAACCAAGGAAGTGA	693
PF		DB	421	GAAGAAACGCGATTGGAAACAGTCGGAATTCCTTCTTACGATTATCAACCAAGGAAGTGA	480
XX	23-JAN-2001; 2001GB-00001760.	QY	694	ACATTTCTGTTTCAAGCCGCTAGTGGAAATTTATACGTTAAAGATGAGGAGGCCCAAGGA	753
PR		DB	481	ACATTTCTGTTTCAAGCCGCTAGTGGAAATTTATACGTTAAAGATGAGGAGGCCCAAGGA	540
XX	(GLAX) GLAXO GROUP LTD.	QY	754	TTTACGCAACAACTTTAAGCCCAATCTAGTGGAACTAGTTGTCCTCCCAACATACGGAATG	813
PI	Edbrooke MR, Lewis AP;	DB	541	TTTACGCAACAACTTTAAGCCCAATCTAGTGGAACTAGTTGTCCTCCCAACATACGGAATG	600
XX	WPI: 2003-150703/15.	QY	814	GATCCAAAATATATGCCCCGCTGATCCAGACTGGATTCCTTTTATACATACGAAATATTT	873
PT	P-PSDB; ASU07720.	DB	601	GATCCAAAATATATGCCCCGCTGATCCAGACTGGATTCCTTTTATACATACGAAATATTT	660
XX	Identifying modulators of serine protease activity useful for treating	QY	874	TGGATATCTAACATCGTAACAGAGAAAGAGAGACTCATTTATGTGCAATAGAGCTA	933
PT	musculoskeletal diseases, by contacting cell expressing a novel serine	DB	661	TGGATATCTAACATCGTAACAGAGAAAGAGAGACTCATTTATGTGCAATAGAGCTA	720
PT	protease polypeptide with a compound and monitoring serine protease	QY	934	GCCAAACATGGAAGAAAGATGCGAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAAGAA	993
PS	Claim 12; Page 22-26; 38pp; English.	DB	721	GCCAAACATGGAAGAAAGATGCGAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAAGAA	780
XX	The invention relates to a method of identifying a substance that	QY	994	TTTGATAGATATCTCGCTATTTGGTGTGTCCTCAAAAGCTGAAACAACTCCCAAGTGGT	1053
CC	modulates serine protease activity, comprising contacting a cell, such as	DB	781	TTTGATAGATATCTCGCTATTTGGTGTGTCCTCAAAAGCTGAAACAACTCCCAAGTGGT	840
CC	a neuronal cell, lung cell, intestinal cell or a cell infected with a	QY	1054	AAAAATCTTACAAATCTTATATGAAGAAATGATGAATCTGAGTGGAAATTTATTCATGTT	1113
CC	virus, expressing a serine protease polypeptide (HIPHUM 46), or its	DB	841	AAAAATCTTACAAATCTTATATGAAGAAATGATGAATCTGAGTGGAAATTTATTCATGTT	900
CC	variant having dipeptidyl peptidase activity, or a serine protease	QY	1114	ACATCCCTATGTTGGAAACAAAGAGGCGAGATTCATTCGTTATCTTAAACACAGGTACA	1173
CC	isolated from the cell with a test substance and monitoring for serine	DB	901	ACATCCCTATGTTGGAAACAAAGAGGCGAGATTCATTCGTTATCTTAAACACAGGTACA	960
CC	protease activity. The method is useful for identifying a substance that	QY	1174	GCAATCTCTAAAGTCATCTTTTAAGATCTCGAAATATGATTCATCTCTGAAGAGAGATC	1233
CC	modulates serine protease activity. A modulator of the serine protease is	DB	961	GCAATCTCTAAAGTCATCTTTTAAGATCTCGAAATATGATTCATCTCTGAAGAGAGATC	1020
CC	useful in the manufacture of a medicament for treatment or prophylaxis of	QY	1234	ATAGATGTCTATAGATAAGAACTAAATTCACCTTTTGGAGATTCCTATTTGAAGGTTGAA	1293
CC	a musculoskeletal disease e.g. osteoarthritis, Hepatitis B virus	DB	1021	ATAGATGTCTATAGATAAGAACTAAATTCACCTTTTGGAGATTCCTATTTGAAGGTTGAA	1080
CC	infection, Alzheimer's disease, paraspinal nerve palsy, myotonic	QY	1294	TATATTTGCCAGAGCTGGATGGAATCTCTGAGGAAATAATGCTTGGTCCATCTCTACTAGAT	1353
CC	dystrophy, Huntington's disease or amyotrophic lateral sclerosis.	DB	1081	TATATTTGCCAGAGCTGGATGGAATCTCTGAGGAAATAATGCTTGGTCCATCTCTACTAGAT	1140
CC	Additional disease that may be treated using modulators of the serine	QY	1354	CGCTCCAGAGCTCGCTTACAGATAGTGTGATCTCACCTGAAATTTATTTATCCAGTAGAA	1413
CC	protease include malabsorption syndromes, irritable bowel syndrome, lung	DB	1141	CGCTCCAGAGCTCGCTTACAGATAGTGTGATCTCACCTGAAATTTATTTATCCAGTAGAA	1200
CC	disease, type I diabetes, faecal incontinence, haemorrhoids, proctitis,	QY	1414	GATGATGTTTATGGAAGGCGAGAGACTCAATGAGTCAGTGCCTGATCTGTGACGCCACTA	1473
CC	rectal polyps, small bowel tumours, colorectal tumours, anaemia,	DB	1201	GATGATGTTTATGGAAGGCGAGAGACTCAATGAGTCAGTGCCTGATCTGTGACGCCACTA	1260
CC	dyslexia, steroid lipofuscinosis, allergic encephalomyelitis, and multiple	QY	1474	ATTATCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTTTCTATGTTTTT	1533
CC	scleriosis. The present sequence represents cDNA encoding the human serine	DB	1261	ATTATCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTTTCTATGTTTTT	1320
CC	protease HIPHUM46 which is located on chromosome 15q21-q22	QY	1534	CCCCAAGTCAGAGAGGAAATTCAGTTATTTTGGCTCTGAACTGCAAAACAGAGTTTC	1593
XX	Sequence 2649 BP; 803 A; 514 C; 585 G; 747 T; 0 U; 0 Other;	DB	1321	CCCCAAGTCAGAGAGGAAATTCAGTTATTTTGGCTCTGAACTGCAAAACAGAGTTTC	1380
XX	Query Match 84.7%; Score 2644.2; DB 8; Length 2649;	QY	1594	CGTCATTTTATACAAAATTCATCTTTTAAAGGAAAGCAATATAAACCGATCCAGTGGT	1653
XX	Best Local Similarity 99.9%; Pred. No. 0;	DB	1381	CGTCATTTTATACAAAATTCATCTTTTAAAGGAAAGCAATATAAACCGATCCAGTGGT	1440
XX	Matches 2646; Conservative 0; Mismatches 3; Indels 0; Gaps 0;	QY	1654	GGGCTGCTGCTCCAAAGTATTTCAAGTGTCTTATCAAGAGGAGTAGCAATTTACCAGT	1713
QY	214 ATGGCAGCAGCAATGGAACACAGACAGCTGGGTGTGAGATATTTGAACTCGGACCTGT	DB	1441	GGGCTGCTGCTCCAAAGTATTTCAAGTGTCTTATCAAGAGGAGTAGCAATTTACCAGT	1500
DB	1 ATGGCAGCAGCAATGGAACACAGACAGCTGGGTGTGAGATATTTGAACTCGGACCTGT	QY	1714	GGTGAATGGGAAGTTCCTTGGCCGGCATGGATCTAATATCCAAAGTTGATGAAGTCAGAGG	1773
QY	274 GAGGAGATATTTGAATCAAGATCGGCTAAATTTGAGCCTTTTATGTTGAGCGGTAT				
DB	61 GAGGAGATATTTGAATCAAGATCGGCTAAATTTGAGCCTTTTATGTTGAGCGGTAT				
QY	334 TCCTGGAGTCAGCTTAAAGCTGCTTCCGATACCGAAATAATCATGCTACATGATG				
DB	121 TCCTGGAGTCAGCTTAAAGCTGCTTCCGATACCGAAATAATCATGCTACATGATG				
QY	394 GCTAAGGCAACCATGATTTTATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTC				
DB	181 GCTAAGGCAACCATGATTTTATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTC				
QY	454 GAGAGATCTATTACCTTCCCATGCTCTGTGAGAACAGAGAAATAACACTGTTTTATTCT				
DB	241 GAGAGATCTATTACCTTCCCATGCTCTGTGAGAACAGAGAAATAACACTGTTTTATTCT				
QY	514 GAAATCCCAAACTATCATAGACAGCAGTCTTAAATGCTCTCTTGGAAAGCCTCTTTTG				
DB	301 GAAATCCCAAACTATCATAGACAGCAGTCTTAAATGCTCTCTTGGAAAGCCTCTTTTG				
QY	574 GATCTTTTTCAGGCAACATGGACTATGGAATGTTATCTCCGAGAGAGAACTATTAAAGA				
DB	361 GATCTTTTTCAGGCAACATGGACTATGGAATGTTATCTCCGAGAGAGAACTATTAAAGA				

Db 1501 GGTGAATGGGAAGTCTTTGGCCGGCATGGATCTAAATATCAAGTTGATGAAGTCAGAAGG 1560
QY 1774 CTGGTATATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACCTGTACGTACGTACGT 1833
Db 1561 CTGGTATATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACCTGTACGTACGTACGT 1620
QY 1834 TACGTAAATCTCGAGAGGTGACCAAGCTGACTGACCGTGGGTACTCAATCTTTGCTGTC 1893
Db 1621 TACGTAAATCTCGAGAGGTGACCAAGCTGACTGACCGTGGGTACTCAATCTTTGCTGTC 1680
QY 1894 ATCAGTCAGCAGCTGCTCTTTTATAGTAAGTATAGTAAGTAAACCAAGAGAAATCCACACTGT 1953
Db 1681 ATCAGTCAGCAGCTGCTCTTTTATAGTAAGTATAGTAAGTAAACCAAGAGAAATCCACACTGT 1740
QY 1954 GTGTCCCTTTTCAAGACTATCAAGTCTCTGAAGATGACCAACTTTGCAAAACCAAGGAATTT 2013
Db 1741 GTGTCCCTTTTCAAGACTATCAAGTCTCTGAAGATGACCAACTTTGCAAAACCAAGGAATTT 1800
QY 2014 TGGGCCACCAATTTTGGATTTCAGCAGGTCTCTTCTGACTATATCTCTCCAGAAATTTTC 2073
Db 1801 TGGGCCACCAATTTTGGATTTCAGCAGGTCTCTTCTGACTATATCTCTCCAGAAATTTTC 1860
QY 2074 TCTTTTGAAGTACTACTGGAATTTACATTTGATTTGGGATGCTCTACAGCCTCATGATCTA 2133
Db 1861 TCTTTTGAAGTACTACTGGAATTTACATTTGATTTGGGATGCTCTACAGCCTCATGATCTA 1920
QY 2134 CAGCTCGAAGAGAAATCTCTACTGCTGCTGTTCATATATGTTGCTCTCAGGTGCAAGTTG 2193
Db 1921 CAGCTCGAAGAGAAATCTCTACTGCTGCTGTTCATATATGTTGCTCTCAGGTGCAAGTTG 1980
QY 2194 GTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCCCTAGCCTCTCTAGT 2253
Db 1981 GTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCCCTAGCCTCTCTAGT 2040
QY 2254 TATGTGGTTGATGATAGACACAGGGGATCTCTGTCACGAGGGCTTAAATTTGAAGGC 2313
Db 2041 TATGTGGTTGATGATAGACACAGGGGATCTCTGTCACGAGGGCTTAAATTTGAAGGC 2100
QY 2314 GCCTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2373
Db 2101 GCCTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2160
QY 2374 CTAGCTCTCCATATGATTTTCATTTGACTTAGATCTGTGGGCATCCACGGCTGGCTAT 2433
Db 2161 CTAGCTCTCCATATGATTTTCATTTGACTTAGATCTGTGGGCATCCACGGCTGGCTAT 2220
QY 2434 GGAGGATACCTCTCCCTGATGGCAATTAATGACAGGTTCAGATATCTTCAGGGTTGCTATT 2493
Db 2221 GGAGGATACCTCTCCCTGATGGCAATTAATGACAGGTTCAGATATCTTCAGGGTTGCTATT 2280
QY 2494 GCTGGGGCCCGTCACTCTGTGATCTTCTATGATACAGATACACGGNACGTTATATG 2553
Db 2281 GCTGGGGCCCGTCACTCTGTGATCTTCTATGATACAGATACACGGNACGTTATATG 2340
QY 2554 GGTCACTCTGACCAATGAACAGGGCTATTACTTAGATCTGTGGCCATGCAAGCAGAA 2613
Db 2341 GGTCACTCTGACCAATGAACAGGGCTATTACTTAGATCTGTGGCCATGCAAGCAGAA 2400
QY 2614 AAGTTCCCTCTGAAACCAAAATCGTTTACTGCTCTTACATGTTTCTCGGATGAGAATGTC 2673
Db 2401 AAGTTCCCTCTGAAACCAAAATCGTTTACTGCTCTTACATGTTTCTCGGATGAGAATGTC 2460
QY 2674 CATTTTGACATACAGATATATTAATGATGTTTTTGTAGGGCTGGAAGCCATATGAT 2733
Db 2461 CATTTTGACATACAGATATATTAATGATGTTTTTGTAGGGCTGGAAGCCATATGAT 2520
QY 2734 TTACAGATCTCTCAGGAGACACAGATATAGTTCTCTGAATCGGAGAACATTAAT 2793
Db 2521 TTACAGATCTCTCAGGAGACACAGATATAGTTCTCTGAATCGGAGAACATTAAT 2580
QY 2794 GAACTGCATCTTTTGCATCTTCAAGAAACCTTGGATCAGGTATTTGCTGCTTAAAA 2853

Db 2581 GAACTGCATCTTTTGCATCTTCAAGAAACCTTGGATCAGGTATTTGCTGCTTAAAA 2640
QY 2854 GTGATATAA 2862
Db 2641 GTGATATAA 2649
RESULT 14
AAH99935
ID AAH99935 standard; cDNA; 2643 BP.
XX
AC AAH99935;
XX
DT 12-APR-2002 (first entry)
XX
Coding sequence of 21953 human prollyl oligopeptidase.
XX
XX 21953 prollyl oligopeptidase; antibody; proline; endopeptidase; cancer;
XX cardiovascular disease; autoimmune disease; atopic allergy;
XX neuronal disorder; vascular disorder; prostate disorder; cytostatic;
XX antidiabetic; antiarthritic; antiasthmatic; antiinflammatory;
XX diabetes mellitus; arthritis; multiple sclerosis; asthma;
XX Grave's disease; neuronal disorder; demyelinating disease; ss.
XX Homo sapiens.
XX
XX WO200179473-A2.
XX
XX 25-OCT-2001.
XX
XX 11-APR-2001; 2001WO-US040483.
XX
XX 18-APR-2000; 2000US-0197508P.
XX (MILL-) MILLENNIUM PHARM INC.
XX
XX Meyers RA, Williamson M;
XX
XX WPI; 2002-034353/04.
XX P-PSDB; AAG78415.
XX
XX New polypeptides 21953, member of human prollyl oligopeptidase family,
XX useful as diagnostic targets and therapeutic agents for controlling
XX cancer, lymphoma and leukemia.
XX
XX Claim 2; Page; 121pp; English.
XX
XX This invention relates to an isolated 21953 human prollyl oligopeptidase.
XX Which is cystostatic, antidiabetic, antiarthritic, neuroprotective,
XX antithyroid, dermatological, antipsoriatic, antiasthmatic,
XX ophthalmological, antiinflammatory, neurotropic, antiparkinsonian,
XX anticonvulsant, gynaecological, vasotropic, antianginal, cardiant,
XX antiatherosclerotic, anorectic and metabolic in its action. Uses include
XX gene therapy, expression or activity of 21953 protein modulator, it is
XX useful for identifying a compound which binds to it and can be used in
XX preventing, treating or detecting a cellular proliferative or
XX differentiative disorder. The 21953 molecules can act as novel diagnostic
XX targets and therapeutic agents for controlling disorders associated with
XX the aberrant activity or degradation of peptide hormones e.g., disorders
XX associated with cell differentiation and proliferation such as cancer,
XX immune function, reproductive, neurological and cardiovascular function.
XX The 21953 molecules are thus useful for treating and preventing cellular
XX proliferative and differentiative disorders, haematopoietic neoplastic
XX disorders, immune disorders such as autoimmune diseases, diabetes
XX mellitus, arthritis, multiple sclerosis, asthma, Grave's disease,
XX neuronal disorders, demyelinating diseases, vascular disorders and
XX metabolism or pain disorders. This sequence represents the cDNA encoding
XX sequence of 21953 human prollyl oligopeptidase. This sequence represents
XX the coding sequence of 21953, being the sequence in between the start and
XX the stop codon of the sequence represented in AAH99934
XX
XX Sequence 2643 BP; 800 A; 514 C; 585 G; 744 T; 0 U; 0 Other;

Query Match									
Best Local Similarity 84.6%; Score 2639.8; DB 6; Length 2643;									
Matches 2641; Conservative 99.9%; Pred. No. 0;									
Mismatches 2; Indels 0; Gaps 0;									
Qy	217	GCAGCAGCAATGGAAAA	CAGAACAGCTGGGTGGTGGAGATATTTGAAACTGGCGACTGTGAG	276					
Db	1	GCAGCAGCAATGGAAAA	CAGAACAGCTGGGTGGTGGAGATATTTGAAACTGGCGACTGTGAG	60					
Qy	277	GAGATATGAATCACAGGATCGGCCTAAATTTGGAGCCCTTTTATGTTGAGCGGTATTCC	336						
Db	61	GAGATATGAATCACAGGATCGGCCTAAATTTGGAGCCCTTTTATGTTGAGCGGTATTCC	120						
Qy	337	TGGAGTCAGCTTAAAGAGCTGCTGCCGATACCAAGAAATATATGAGTATGATGAGTGGCT	396						
Db	121	TGGAGTCAGCTTAAAGAGCTGCTGCCGATACCAAGAAATATATGAGTATGATGAGTGGCT	180						
Qy	397	AAGGCACACATGATTTTCATGTTTGAAGAGGATGATCCAGATGGAGCTCATTCAGAC	456						
Db	181	AAGGCACACATGATTTTCATGTTTGAAGAGGATGATCCAGATGGAGCTCATTCAGAC	240						
Qy	457	AGAAATCTAATACCTTGGCATGTCTGTGTGAGAACAGAGAAATACACTGTTTATTCGAA	516						
Db	241	AGAAATCTAATACCTTGGCATGTCTGTGTGAGAACAGAGAAATACACTGTTTATTCGAA	300						
Qy	517	ATTCCCAAAAATATCAATAGAGCAGCAGTCTTAATGCTCTCTTGGAGGCTCTTTTGGAT	576						
Db	301	ATTCCCAAAAATATCAATAGAGCAGCAGTCTTAATGCTCTCTTGGAGGCTCTTTTGGAT	360						
Qy	577	CTTTTTCAGGCAACACTGGACTATGGAATGATTTCTCGAGAGAGAACTATTAAGAGAA	636						
Db	361	CTTTTTCAGGCAACACTGGACTATGGAATGATTTCTCGAGAGAGAACTATTAAGAGAA	420						
Qy	637	AGAAACCGATGGAAACAGTCGGAATTCCTTACAGTATATCACCAGGAGTGGAA	696						
Db	421	AGAAACCGATGGAAACAGTCGGAATTCCTTACAGTATATCACCAGGAGTGGAA	480						
Qy	697	TTTTCTGTTTCAAGCCGTAGTGGAAATTTATCAGTAAAGATGGAGGCCACAGGATTT	756						
Db	481	TTTTCTGTTTCAAGCCGTAGTGGAAATTTATCAGTAAAGATGGAGGCCACAGGATTT	540						
Qy	757	ACGCAACAACTTTAAGGCCCAATCTATAGTGGAACTAGTTGTCTCCCAACATACGGATGAT	816						
Db	541	ACGCAACAACTTTAAGGCCCAATCTATAGTGGAACTAGTTGTCTCCCAACATACGGATGAT	600						
Qy	817	CCAAAATTTATGCCCGCTGTATCCAGACTGGATTGCTTTTATACATAGCAACGATATTTGG	876						
Db	601	CCAAAATTTATGCCCGCTGTATCCAGACTGGATTGCTTTTATACATAGCAACGATATTTGG	660						
Qy	877	ATATCTAACATCGTAACCCAGAGAAAGAGAGACTCACTTATGTGCACAATGAGCTAGCC	936						
Db	661	ATATCTAACATCGTAACCCAGAGAAAGAGAGACTCACTTATGTGCACAATGAGCTAGCC	720						
Qy	937	AACATGGAAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTTCCAAAGAAATTT	996						
Db	721	AACATGGAAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTTCCAAAGAAATTT	780						
Qy	997	GATAGATATTCGGCTATGTTGGTGTCCAAAGCTGAAACACTCCAGTGGTGTAAA	1056						
Db	781	GATAGATATTCGGCTATGTTGGTGTCCAAAGCTGAAACACTCCAGTGGTGTAAA	840						
Qy	1057	ATTCCTAGAAATCTATATGAAGAAATGATGATCTGAGTGGAAATTTTCACTGTATCA	1116						
Db	841	ATTCCTAGAAATCTATATGAAGAAATGATGATCTGAGTGGAAATTTTCACTGTATCA	900						
Qy	1117	TCCCTATGTTGGAACCAAGAGGGCAGATTCATTCGGTTATCCTAAAACAGGTACAGCA	1176						
Db	901	TCCCTATGTTGGAACCAAGAGGGCAGATTCATTCGGTTATCCTAAAACAGGTACAGCA	960						
Qy	1177	AATCCTAAAGTCACCTTTTAAAGATGTCAGAAATATATGATGCTGAAGGAAAGGATCAT	1236						
Db	961	AATCCTAAAGTCACCTTTTAAAGATGTCAGAAATATATGATGCTGAAGGAAAGGATCAT	1020						
Qy	1237	GATGTATAGATAAGGAACCTAAATCAACCTTTTGGATTTCTATTTGAAGGAGTTGAATAT	1296						

1021	GATGTATAGATAAGGAACCTAAATCAACCTTTTGGATTTCTATTTGAAGGAGTTGAATAT	1080
1297	ATTGCGAGAGCTGGATGGACTCCTGAGGAAAAATATGCTTTGGTCCATCCCTACTAGATCGC	1356
1081	ATTGCGAGAGCTGGATGGACTCCTGAGGAAAAATATGCTTTGGTCCATCCCTACTAGATCGC	1140
1357	TCCAGAGCTCGCCTACAGATAGTGTGATCTCACCTGAATATTTATTTCCAGGTAGAGAT	1416
1141	TCCAGAGCTCGCCTGCAGATAGTGTGATCTCACCTGAATATTTATTTCCAGGTAGAGAT	1200
1417	GATGTTATGGAAGGAGAGACTCATTTGAGTCAGTGGCTGATCTGTGAGCCCACTAAAT	1476
1201	GATGTTATGGAAGGAGAGACTCATTTGAGTCAGTGGCTGATCTGTGAGCCCACTAAAT	1260
1477	ATCTATGAGAAACAAACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTTCCC	1536
1261	ATCTATGAGAAACAAACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTTCCC	1320
1537	CAAGTCAGGAGAGGAAATTTGAGTTTATTTTGGCTCTGATGCAAAAACAGGTTCCGT	1596
1321	CAAGTCAGGAGAGGAAATTTGAGTTTATTTTGGCTCTGATGCAAAAACAGGTTCCGT	1380
1597	CATTTATCAAAATTTACATCTATTTTAAAGGAAAGCAAAATATAAAGATCCAGTGGTGG	1656
1381	CATTTATCAAAATTTACATCTATTTTAAAGGAAAGCAAAATATAAAGATCCAGTGGTGG	1440
1657	CTGCTGCTCCAAAGTGATTTCAAGTGCTCTATCAAAAGGAGATAGCAATTTACAGTGGT	1716
1441	CTGCTGCTCCAAAGTGATTTCAAGTGCTCTATCAAAAGGAGATAGCAATTTACAGTGGT	1500
1717	GAATGGAAAGTTCTTTGGCCGGCATGGATCTAAATATCAAGTTGATGAAGTCAGAAAGCTG	1776
1501	GAATGGAAAGTTCTTTGGCCGGCATGGATCTAAATATCAAGTTGATGAAGTCAGAAAGCTG	1560
1777	GTATATTTTGAAGGACCAAAAGACTCCCTTTTAGAGCATCACCTGTACGTAGTCAGTTAC	1836
1561	GTATATTTTGAAGGACCAAAAGACTCCCTTTTAGAGCATCACCTGTACGTAGTCAGTTAC	1620
1837	GTAAATCTCGAGAGGTCAGAGGCTGACCGCTGACTGACCGTGGCTACTCACATTTCTTGTGATC	1896
1621	GTAAATCTCGAGAGGTCAGAGGCTGACCGCTGACTGACCGTGGCTACTCACATTTCTTGTGATC	1680
1897	AGTCAGACTGTGACTTCTTTATAGTAAGTATAGTAACCCAGAGAGATCCACACTGTGTG	1956
1681	AGTCAGACTGTGACTTCTTTATAGTAAGTATAGTAACCCAGAGAGATCCACACTGTGTG	1740
1957	TCCCTTTACAGCTATCAAGTCTCTGAGATGACCTGCAAACTTGCRAAAACAAAGGAAATTTGG	2016
1741	TCCCTTTACAGCTATCAAGTCTCTGAGATGACCTGCAAACTTGCRAAAACAAAGGAAATTTGG	1800
2017	GCCACCAATTTGGATTCAGCAGGTCCTCTTCTGACTATATCTCTCAGAAAATTTTCTCT	2076
1801	GCCACCAATTTGGATTCAGCAGGTCCTCTTCTGACTATATCTCTCAGAAAATTTTCTCT	1860
2077	TTTGAAGTACTACTGGATTTTACATTTGATGGGATGCTCTACAAGCCTCATGATCTACAG	2136
1861	TTTGAAGTACTACTGGATTTTACATTTGATGGGATGCTCTACAAGCCTCATGATCTACAG	1920
2137	CCTGGAAGAAATATCTACTGTGCTTTTATATATGTTGTTGCTCTCAGGTGTCAGTTGGT	2196
1921	CCTGGAAGAAATATCTACTGTGCTTTTATATATGTTGTTGCTCTCAGGTGTCAGTTGGT	1980
2197	AATAATCCGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTTAGCTCTCTAGGTTAT	2256
1981	AATAATCCGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTTAGCTCTCTAGGTTAT	2040
2257	GTGGTTGTAGTATAGCAAGAGGGATCTGTGACCGAGGGCTTAATTTGAAGGCGCC	2316
2041	GTGGTTGTAGTATAGCAAGAGGGATCTGTGACCGAGGGCTTAATTTGAAGGCGCC	2100
2317	TTTAAATATAAATGGGTCAAAATAGAAATTTGACGATCAGGTGGAGAGACTCCATATCTA	2376

Db 2101 TTTAAATATAAATGGGTCAAAATAGAAATGAGATCAGGTGGAAGGACTCAATATCTA 2160
Qy 2377 GCTTCTCGATATGATTTTCAATGACTTAGATCGTGTGGGCATCCACGGCTGGTCTTATGGA 2436
Db 2161 GCTTCTCGATATGATTTTCAATGACTTAGATCGTGTGGGCATCCACGGCTGGTCTTATGGA 2220
Qy 2437 GGATACCTCTCCCTGATGAGCAATTAATGCGAGAGTCAGATATCTTACGGGTTGCTATTGCT 2496
Db 2221 GGATACCTCTCCCTGATGAGCAATTAATGCGAGAGTCAGATATCTTACGGGTTGCTATTGCT 2280
Qy 2497 GGGGCCCCAGTCACTCTGTGATCTTCTATGATACAGGATACACGGAACGTTTATATGGGT 2556
Db 2281 GGGGCCCCAGTCACTCTGTGATCTTCTATGATACAGGATACACGGAACGTTTATATGGGT 2340
Qy 2557 CACCTGCACAGAAATGAACAGGCTATTTACTTAGGATCTGTGGCCATGCAAGCAGAAAAG 2616
Db 2341 CACCTGCACAGAAATGAACAGGCTATTTACTTAGGATCTGTGGCCATGCAAGCAGAAAAG 2400
Qy 2617 TTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGTTTCTCTGGATGAGAAATGTCAT 2676
Db 2401 TTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGTTTCTCTGGATGAGAAATGTCAT 2460
Qy 2677 TTTCACATACCAATATTTACTGAGTCTTTTCTAGTGGCTGGAAGCCATATGATTTA 2736
Db 2461 TTTCACATACCAATATTTACTGAGTCTTTTCTAGTGGCTGGAAGCCATATGATTTA 2520
Qy 2737 CAGATCTATCTCAGGAGAGACACAGCATAGAGTTTCTGAATCGGAGAAACATTAATGAA 2796
Db 2521 CAGATCTATCTCAGGAGAGACACAGCATAGAGTTTCTGAATCGGAGAAACATTAATGAA 2580
Qy 2797 CTGATCTTTTGGACTACCTTTCAAGAAACCTTTGGATCAGTATTGCTGCTCTAAAAGTG 2856
Db 2581 CTGATCTTTTGGACTACCTTTCAAGAAACCTTTGGATCAGTATTGCTGCTCTAAAAGTG 2640
Qy 2857 ATA 2859
Db 2641 ATA 2643

RESULT 15
ABK83325
ID ABK83325 standard; cDNA; 4523 BP.
XX AC ABK83325;
XX DT 12-AUG-2002 (first entry)
XX DE cDNA encoding human DPRP-1 splice variant #1.
XX KW Human; serine protease; dipeptidyl peptidase IV-related protein; DPRP;
KW DPPIV; infection; human immunodeficiency virus; HIV-1; HIV-2; pain;
KW diabetes; infertility; obesity; anorexia; Parkinson's disease; stroke;
KW heart failure; hypertension; urinary retention; osteoporosis; cancer;
KW ulcer; allergy; cancer; psychotic disorder; neurological disorder;
KW dyskinesia; reproductive disorder; inflammatory disorder;
KW metabolic disorder; gene; ss.
XX OS Homo sapiens.
XX PN WO200231134-A2.
XX PD 18-APR-2002.
XX PF 12-OCT-2001; 2001WO-US031874.
XX PR 12-OCT-2000; 2000US-0240117P.
XX (FERR) FERRING BV.
XX Q1 S, Akinsanya KO, Riviere PJ, Junien J;
XX WPI: 2002-444178/47.
XX P-PSDB; ABG61594.

XX PT New dipeptidyl peptidase IV-related proteins and nucleic acids encoding
PT the proteins, useful for treating e.g. fungal, bacterial, protozoan and
PT viral infections, cancers, allergies, neurological disorders, or pain.
XX PS Disclosure; Page 61-62; 113pp; English.
XX CC The present invention relates to the isolation of novel human serine
CC proteases referred to as dipeptidyl peptidase IV (DPPIV)-related proteins
CC (DPRP). The dipeptidyl peptidase IV-related proteins (DPRP) and nucleic
CC acids encoding them are useful for treating infections such as fungal,
CC bacterial, protozoan and viral infections, particularly infections caused
CC by human immunodeficiency virus (HIV-1 or HIV-2), pain, diabetes,
CC precocious puberty, infertility, obesity, anorexia, bulimia, Parkinson's
CC disease, acute heart failure, hypotension, hypertension, urinary
CC retention, osteoporosis, angina pectoris, stroke, ulcers, asthma,
CC allergies, cancers, migraine, vomiting, psychotic and neurological
CC disorders (e.g. anxiety, dementia, or schizophrenia), and dyskinesias.
CC These may also be used in discovering therapeutic agents for the
CC treatment of reproductive, inflammatory and metabolic disorders. ABK83322
CC -ABK83343 encode human DPRP proteins
XX SQ Sequence 4523 BP; 1384 A; 828 C; 940 G; 1371 T; 0 U; 0 Other;
Query Match 80.1%; Score 2500.4; DB 6; Length 4523;
Best Local Similarity 90.2%; Pred. No. 0;
Matches 2815; Conservative 0; Mismatches 1; Indels 304; Gaps 1;
Qy 1 AAGTGCTAAAGCCTCCGAGCAAGCGCGTGTCTACTGCGCGCGCTCTTTAGTGGCG 60
Db 1 AAGTGCTAAAGCCTCCGAGCAAGCGCGTGTCTACTGCGCGCGCTCTTTAGTGGCG 60
Qy 61 CGTTGCCCGCTGGGTTTGTCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
Db 61 CGTTGCCCGCTGGGTTTGTCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
Qy 121 GAGTGAGCG 180
Db 121 GAGTGAGCG 180
Qy 181 TCG 240
Db 181 TCG 240
Qy 241 CTGGGTGTTGAGATATTTGAAATCTCGGACTGTGAGGAGATATTTGAATCAGGATCGG 300
Db 241 CTGGGTGTTGAGATATTTGAAATCTCGGACTGTGAGGAGATATTTGAATCAGGATCGG 300
Qy 301 CTTAAATTTGAGCGCTTTTATGTTGAGCGGTATTCCTGGAGTCACTTAAAAAGCTGCTT 360
Db 301 CTTAAATTTGAGCGCTTTTATGTTGAGCGGTATTCCTGGAGTCACTTAAAAAGCTGCTT 360
Qy 361 GCGATACAGAAATATCCTGGCTACATGATGCTAGGCGCGCGCGCGCGCGCGCGCGCG 420
Db 361 GCGATACAGAAATATCCTGGCTACATGATGCTAGGCGCGCGCGCGCGCGCGCGCGCG 420
Qy 421 GTGAAGAGGATGATCCAGATGACCTCATTCAGACAGATCTATTACCTTGGCATGCTCT 480
Db 421 GTGAAGAGGATGATCCAGATGACCTCATTCAGACAGATCTATTACCTTGGCATGCTCT 480
Qy 481 GGTGAGACAGAGAAAATACACTGTTTATTTGAAATTCCTGAAATTCCTGAAATTCCTG 540
Db 481 GGTGAGACAGAGAAAATACACTGTTTATTTGAAATTCCTGAAATTCCTGAAATTCCTG 540
Qy 541 GCAGTCTTAATGCTCTCTTTGGAAGCTCTTTTGGATCTTTTTCAGGCGCACTGGACTAT 600
Db 541 GCAGTCTTAATGCTCTCTTTGGAAGCTCTTTTGGATCTTTTTCAGGCGCACTGGACTAT 600
Qy 601 GGAATGATTTCTCGAGAGAGAACTATTTAAGAGAGAGAAAGCGCATTTGGAACAGTGGGA 660
Db 601 GGAATGATTTCTCGAGAGAGAACTATTTAAGAGAGAGAAAGCGCATTTGGAACAGTGGGA 660
Qy 661 ATTGCTTCTTACGATTTATCAACAGAGAGTGGAACTTTCTGTTTCAAGCGCGTAGTGGGA 720

Qy	2881	TCTCTGGTATACACTGGCTATTAAACCAAAATGAGGAGGTTTAATCAACAGAAAAACACAGA	2940
Db	2577	TCTCTGGTATACACTGGCTATTAAACCAAAATGAGGAGGTTTAATCAACAGAAAAACACAGA	2636
Qy	2941	ATTGATCATCACATTTTGATACCTGCCATGTAAACATCTACTCTCTGAAAAATAAATGTGGTG	3000
Db	2637	ATTGATCATCACATTTTGATACCTGCCATGTAAACATCTACTCTCTGAAAAATAAATGTGGTG	2696
Qy	3001	CCATGCAGGGGCTACGGTTTGTGGTGTAGTAAATCTAATACCTTTAACCCACATGCTCAAAA	3060
Db	2697	CCATGCAGGGGCTACGGTTTGTGGTGTAGTAAATCTAATACCTTTAACCCACATGCTCAAAA	2756
Qy	3061	TCAAATCATACATATTCTCTGAGAGACCCAGCAATACCATAGAAATTACTAAAAA	3120
Db	2757	TCAAATCATACATATTCTCTGAGAGACCCAGCAATACCATAGAAATTACTAAAAA	2816

Search completed: April 15, 2006, 00:48:48
Job time : 1779 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 14, 2006, 19:46:05 ; Search time 10952 Seconds
(without alignments)
13328.679 Million cell updates/sec

Title: US-10-825-632-2
Perfect score: 3120
Sequence: 1 aagtgttaagctccgagg.....agaattactaaaaaaa 3120

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0
Searched: 41078325 segs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_btc:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gsa1:*
10: gb_gsa2:*
11: gb_gsa3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2446.8	78.4	2649	10 AY411615	AY411615 Homo sapi
2	2381.4	76.3	2649	10 AY411616	AY411616 Pan trogl
3	2251.8	72.2	3143	4 AK016546	AK016546 Mus muscu
4	2158.6	69.2	5517	4 AK029788	AK029788 Mus muscu
5	1985	63.6	2634	10 AY411617	AY411617 Mus muscu
6	1948.8	62.5	2292	4 CR609512	CR609512 full-leng
7	1217.4	39.0	1292	4 AF175225	AF175225 Homo sapi
8	1200	38.5	1265	4 AF176779	AF176779 Homo sapi
9	1132	36.3	1278	4 AF173382	AF173382 Homo sapi
10	938.2	30.1	1042	3 BM557438	BM557438 AGENCOURT
11	882.2	28.3	3327	4 AK050021	AK050021 Mus muscu
12	880.6	28.2	3457	4 AK078301	AK078301 Mus muscu
13	875.8	28.1	884	7 CR985650	CR985650 CR985650
14	871.2	27.9	3376	4 AK050023	AK050023 Mus muscu
15	848.6	27.2	957	5 BQ675006	BQ675006 AGENCOURT
16	848.2	27.2	881	7 CR990494	CR990494 CR990494
17	845.6	27.1	910	5 BQ675260	BQ675260 AGENCOURT
18	844.4	27.1	910	5 BX372276	BX372276 BX372276
19	837.4	26.8	985	3 BQ068650	BQ068650 AGENCOURT
20	827	26.5	864	1 AL542617	AL542617 AL542617
21	818	26.2	1013	7 CN647232	CN647232 ILLUMIGEN
22	818	26.2	1027	7 CN646612	CN646612 ILLUMIGEN

c	23	812.2	26.0	910	2	BG479035	BG479035 602526111
	24	804.4	25.8	902	8	DN517062	DN517062 1257572 M
	25	803.6	25.8	914	5	BX390898	BX390898 BX390898
	26	801.8	25.7	835	7	CR998849	CR998849 CR998849
	27	801.2	25.7	1041	2	BI084090	BI084090 602869453
	28	784	25.1	968	5	BQ671635	BQ671635 AGENCOURT
c	29	775.4	24.9	855	2	BI084885	BI084885 602869453
	30	768	24.6	857	8	DN519771	DN519771 1262448 M
c	31	763.6	24.5	804	5	BG687687	BG687687 UI-CF-EC1
	32	761.6	24.4	789	2	BG709118	BG709118 602675382
	33	758.4	24.3	840	8	DN514280	DN514280 1251358 M
	34	756.6	24.3	820	7	CR994557	CR994557 CR994557
	35	750.4	24.1	753	1	AL040398	AL040398 DKF2p434A
	36	742.4	23.8	853	2	BI223892	BI223892 602941035
	37	736.4	23.6	742	7	CN427229	CN427229 170006002
	38	734	23.5	735	1	AL043338	AL043338 DKF2p434O
c	39	726.4	23.3	760	1	AI917735	AI917735 tt11d07.x
	40	719.6	23.1	826	3	BI553405	BI553405 603193351
	41	706.2	22.6	1002	2	BE733691	BE733691 601569288
	42	697.8	22.4	726	7	CN427231	CN427231 170004554
	43	694	22.2	724	7	CK000768	CK000768 AGENCOURT
	44	691.6	22.2	788	2	BG163397	BG163397 602338360
c	45	688.8	22.1	708	5	BU622228	BU622228 UI-H-FH1-

ALIGNMENTS

RESULT 1	AY411615	2649 bp	DNA	linear	GSS 16-DEC-2003
AY411615	LOCUS	Homo sapiens DPP8 gene, genomic survey sequence.	VIRTUAL TRANSCRIPT, partial sequence,		
DEFINITION	AY411615	GI:39767583			
ACCESSION	AY411615				
VERSION	AY411615.1				
KEYWORDS	GSS.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1 (bases 1 to 2649)				
AUTHORS	Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarimal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferreria, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J., Adams, M.D. and Cargill, M.				
TITLE	Inferred nonneutral evolution from human-chimp-mouse orthologous gene trios				
JOURNAL	Science 302 (5652), 1960-1963 (2003)				
PUBMED	14671302				
REFERENCE	2 (bases 1 to 2649)				
AUTHORS	Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarimal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferreria, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J., Adams, M.D. and Cargill, M.				
TITLE	Direct Submision				
JOURNAL	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA				
COMMENT	This sequence was made by sequencing genomic exons and ordering them based on alignment.				
FEATURES	Location/Qualifiers				
source	1..2649				
	/organism="Homo sapiens"				
	/mol_type="genomic DNA"				
gene	/db_xref="taxon:9606"				
	<1..2649				
	/gene="DPP8"				
	/locus_tag="HCM4283"				
ORIGIN					
Query Match	78.4%	Score	2446.8	DB	10
Best Local Similarity	92.4%	Pred. No.	0	Length	2649
Matches	2448	Conservative	0	Mismatches	201
		Indels	0	Gaps	0


```
|||||
Db 2161 CTAGCTTCGATATGATTTCAATGCTTAGATCGTGTGGGCATCCACGGCTGGTCCAT 2220
|||||
Qy 2434 GGAGGATACCTCTCCCTGATGGCATTAATGACAGAGTCAGATATCTTTCAGGGTGTCTATT 2493
|||||
Db 2221 GGAGGATACCTCTCCCTGATGGCATTAATGACAGAGTCAGATATCTTTCAGGGTGTCTATT 2280
|||||
Qy 2494 GCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACTGTATATG 2553
|||||
Db 2281 GCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACTGTATATG 2340
|||||
Qy 2554 GGTACCCCTGACAGAAATGAACAGGGCTATTACTTTAGGATCTGTGGCCATGCAAGCANN 2613
|||||
Db 2341 GGTACCCCTGACAGAAATGAACAGGGCTATTACTTTAGGATCTGTGGCCATGCAAGCANN 2400
|||||
Qy 2614 AAGTTCCTCTGACCAATCGTTTACTGCTCTTACATGTTTCTTGGATGAGAATGTC 2673
|||||
Db 2401 NNNNNNNNNNNNACAAATCGTTTACTGCTCTTACATGTTTCTTGGATGAGAATGTC 2460
|||||
Qy 2674 CATTTTGCACATACCACTATATTAATGAGTGTGAGGGCTGGAAAGCCATATGAT 2733
|||||
Db 2461 CATTTTGCACATACCACTATATTAATGAGTGTGAGGGCTGGAAAGCCATATGAT 2520
|||||
Qy 2734 TTACAGATCTATCCTCAGAGAGACACAGCATAAGAGTTCCTGAATCGGGAGAACATTAT 2793
|||||
Db 2521 TTACAGATCTATCCTCAGAGAGACACAGCATAAGAGTTCCTGAATCTGGAGAACATTAT 2580
|||||
Qy 2794 GAATCGATCTTTTGCACCTACCTTCAAGAAACCTTGGATCACCTATGCTGCTCTAAAA 2853
|||||
Db 2581 GAATCGATCTTTTGCACCTACCTTCAAGAAACCTTGGATCACCTATGCTGCTCTAAAA 2640
|||||
Qy 2854 GTGATATAA 2862
|||||
Db 2641 GTGATATAA 2649

RESULT 2
AV411616 2649 bp DNA linear GSS 16-DEC-2003
LOCUS Pan troglodytes DPP8 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AV411616
VERSION AV411616.1 GI:39767584
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Pan.
1 (bases 1 to 2649)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
Todd,M.A., Tanenbaum,D.M., Civeillo,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
14671302
2 (bases 1 to 2649)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
Todd,M.A., Tanenbaum,D.M., Civeillo,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment.
Location/Qualifiers
1 .2649
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"

FEATURES
source
```

```
gene <1..>2649
/genes="DPP8"
/locus_tag="HCM4283"

ORIGIN

Query Match 76.3%; Score 2381.4; DB 10; Length 2649;
Best Local Similarity 90.3%; Pred. No. 0;
Matches 2388; Conservative 0; Mismatches 261; Indels 0; Gaps 0;

Qy 214 ATGCGACAGCAATGGAAACAGAAACAGCTGGGTGTGGAGATATTTGAAACCTGCGACTGT 273
|||||
Db 1 ATGCGACAGCAATGGAAACAGAAACAGCTGGGTGTGGAGATATTTGAAACCTGCGACTGT 60
|||||
Qy 274 GAGGAGAAATTTGAATCACAGGATCGGCTAAATTTGAGGCCCTTTTATGTTGAGCGGTAT 333
|||||
Db 61 GAGGAGAAATTTGAATCACAGGATCGGCTAAATTTGAGGCCCTTTTATGTTGAGCGGTAT 120
|||||
Qy 334 TCCTGGAGTCAGCTTTAAAGCTGCTTGGCGATACCAGAAATATCATGGCTACATGATG 393
|||||
Db 121 TCCTGGAGTCAGCTTTAAAGCTGCTTGGCGATACCAGAAATATCATGGCTACATGATG 180
|||||
Qy 394 GCTAAGGCACACATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTTCA 453
|||||
Db 181 GCTAAGGCACACATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGGGCTCATTTCA 240
|||||
Qy 454 GACAGAAATCTATTACCTTGGCCATGCTCTGTGTGAGAACAGAGAAATATACACTGTTTATTCT 513
|||||
Db 241 GACAGAAATCTATTACCTTGNNNNNNGGTGAGAACAGAGAAATATACACTGTTTATTCT 300
|||||
Qy 514 GAAATTCCTAAAACATATCAATAGACAGCAGCTTTAATGCTCTCTTGGAGGCTCTTTTG 573
|||||
Db 301 GAAATTCCTAAAACATATCAATAGACAGCAGCTTTAATGCTCTCTTGGAGGCTCTTTTG 360
|||||
Qy 574 GATCTTTTTCAGGCAACACCTGGACTATGGAATGTATTCTCGAGAGAGAAAGAACTATTAAAG 633
|||||
Db 361 GATCTTTTTCAGGCAACACCTGGACTATGGAATGTATTCTCGAGAGAGAAAGAACTATTAAAG 420
|||||
Qy 634 GAAAGAAACCGCATTTGSAACAGTCGGAATTTGCTTTTACGATTTATCACAAGGAAGTCGA 693
|||||
Db 421 GAAAGAAACCGCATTTGSAACAGTCGGAATTTGCTTTTACGATTTATCACAAGGAAGTCGA 480
|||||
Qy 694 ACATTTCTGTTTTCAGCCGCTAGTGAATTTATCAGTAAAGATGGAGGGCCACAAGGA 753
|||||
Db 481 ACATTTCTGTTTTCAGCCGCTAGTGAATTTATCAGTAAAGATGGAGGGCCACAAGGA 540
|||||
Qy 754 TTTACGCAACAACTTTTAAAGCCCAATCTAGTGGAACTAGTGTGCCCAACATACGGATG 813
|||||
Db 541 TTTACGCAACAACTTTTAAAGCCCAATCTAGTGGAACTAGTGTGCCCAACATACGGATG 600
|||||
Qy 814 GATCCAAAATTTATGCCCGCTGATCCAGACTGGATTTGCTTTTATACATAGCAACGATATT 873
|||||
Db 601 GATCCAAAATTTATGCCCGCTGATCCAGACTGGATTTGCTTTTATACATAGCAACGATATT 660
|||||
Qy 874 TGGATATCTAACATCGTAAACAGAGAAAGAGAGACTCACTTATGTGTCACATGAGCTA 933
|||||
Db 661 TGGATATCTAACATCGTAAACAGAGAAAGAGAGACTCACTTATGTGTCACATGAGCTA 720
|||||
Qy 934 GCCAACATGSAAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCCAAGAA 993
|||||
Db 721 GCCAACATGSAAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCCAAGAA 780
|||||
Qy 994 TTTGATAGATTTCTGGCTATTGCTGGTGTGTCGCAAGAGCTGAAACAACTCCCAAGTGGT 1053
|||||
Db 781 NNNNATAGATATNCTGGCTATTGCTGGTGTGTCGCAAGAGCTGAAACAACTCCCAAGTGGT 840
|||||
Qy 1054 AAAATTTCTAGAAATTTCTATATGAAGAAATATGATGAAATCTGAGGTGAAATTTATGTT 1113
|||||
Db 841 AAAATTTCTAGAAATTTCTATATGAAGAAATATGATGAAATCTGAGGTGAAATTTATGTT 900
|||||
Qy 1114 ACATCCCTGATTTGGAAACAGAGGGGCGAGATTCTATCCGTTATCCTTAAACAGGTACA 1173
|||||
Db 901 ACATCCCTGATTTGGAAACAGAGGGGCGAGATTCTATCCGTTATCCTTAAACAGGTACA 960
|||||
```

Qy	1174	GCAATCCTAAAGTCACTTTTAAGATGTCAGAAATATATGATGATGCTGAAGAGGATC	1233
Db	961	GCAATNCNTAAAGTCACTTTTAAGATGTCAGAAATATATGATGATGCTGAAGAGGANN	1020
Qy	1234	ATAGATGTCATAGATAAGAGAACTAAATCAACCTTTTGGAGATTCATTTGAAGAGATTGAA	1293
Db	1021	NN	1080
Qy	1294	TATATTGCCAGAGCTGGATGGACTCCTGAGGAAATATATCTTTGGTCCATCTCTAGAT	1353
Db	1081	NN	1140
Qy	1354	CGCTCCGAGCTCGCTACAGATAGTTGATCTCACTGAAATATTTATCCAGTAGAA	1413
Db	1141	CGCTCCGAGCTCGCTCGAGATAGTTGATCTCACTGAAATATTTATCCAGTAGAA	1200
Qy	1414	GATGATGTTATGGAAGGACAGAGCTCATTTGAGTCAGTGCCTGATTTCTGTGAGCCACTA	1473
Db	1201	GATGATGTTATGGAAGGACAGAGCTCATTTGAGTCAGTGCCTGATTTCTGTGAGCCACTA	1260
Qy	1474	ATTATCTATGAAGAAACAAAGACATCTGGATAAATATCCATGACATCTTCATGTTTTT	1533
Db	1261	ATTANCATGAAGAAACAAAGACATCTGGATAAATATCCATGACATCTTCATGTTTTT	1320
Qy	1534	CCCCAAAGTCAAGAGAGGAAATTGAGTTTATTTTTCCTCTGATTCGAAACAGGTTTC	1593
Db	1321	CCCCAAAGTCAAGAGAGGAAATTGAGTTTATTTTTCCTCTGATTCGAAACAGGTTTC	1380
Qy	1594	CGTCATTTATCAAAATTTACATCTATTTTAAAGGAAACAAATATAAAGATCCAGTGT	1653
Db	1381	CGTCATTTATCAAAATTTACATCTATTTTAAAGGAAACAAATATAAAGATCCAGTGT	1440
Qy	1654	GGCTCGCTGCTCCAAAGTGATTTCAAGTGTCTATCAAGAGGAGATAGCAATTTACCAGT	1713
Db	1441	GGCTCGCTGCTCCAAAGTGATTTCAAGTGTCTATCAAGAGGAGATAGCAATTTACCAGT	1500
Qy	1714	GGTGAATGGGAAGTTCTTGGCGCGCATGGATCTAATATCCAAAGTTGATGAAGTCAGAAG	1773
Db	1501	NN	1560
Qy	1774	CTCGTATATTTGAAGGACCAAGACTCCCTTTTAGAGCATCACCTGTACGTAGTCAGT	1833
Db	1561	CTCGTATATTTGAAGGACCAAGACTCCCTTTTAGAGCATCACCTGTACGTAGTCAGT	1620
Qy	1834	TAGTAAATCTCGAGAGTGA CAAGGCTGACTGACCGTGCTACTCACATTTCTTGCTGC	1893
Db	1621	TAGTAAATCTCGAGAGTGA CAAGGCTGACTGACCGTGCTACTCACATTTCTTGCTGC	1680
Qy	1894	ATCAGTCAGCACTGTGACTCTTTATAGTAAGTATAGTACCAAGAGAAATCCACTGT	1953
Db	1681	ATCAGTCAGCACTGTGACTCTTTATAGTAAGTATAGTACCAAGAGAAATCCACTGT	1740
Qy	1954	GTGTCCCTTTA CAAGCTATCAAGTCTGAAGATGACCAACTTGCAGAAACAAAGGAATTT	2013
Db	1741	GTGTCCCTTTA CAAGCTATCAAGTCTGAAGATGACCAACTTGCAGAAACAAAGGAATTT	1800
Qy	2014	TGGGCCACCACTTTTGGATTCAGCAGGTCTCTTCTGACTATATCTCTCCAGAAATTTTC	2073
Db	1801	TGGGCCACCACTTTTGGATTCAGCAGGTCTCTTCTGACTATATCTCTCCAGAAATTTTC	1860
Qy	2074	TCCTTTGAAAGTACTACTGATTTACATTTGATGGATGCTCTCAAGCCCTCATGATCTA	2133
Db	1861	TCCTTTGAAAGTACTACTGATTTACATTTGATGGATGCTCTCAAGCCCTCATGATCTA	1920
Qy	2134	CAGCTCGAAAGAAATATCTACTGTCTGTTTCAATATGTTGGTGGTCTCCAGGTGCAAGTTG	2193
Db	1921	CAGCTCGAAAGAAATATCTACTGTCTGTTTCAATATGTTGGTGGTCTCCAGGTGCAAGTTG	1980
Qy	2194	GTGAATATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCCCTGCTCTAGGT	2253
Db	1981	GTGAATATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCCCTGCTCTAGGT	2040
Qy	2254	TATGTGTTGTAGTAGACACAGGGATCCTGTCCCGAGGCTTAAATTTTGAAGGC	2313

Db	2041	TAITGGTTTAGTGATAGA CAA CAGGGATCCTGTCCACGAGGCTTAAATTTGAAGGC	2100
Qy	2314	GCTTTAAATATAAATGGGTCAAAATAGAAATTTGACGATCAGGTGGAAGACTCCAAATAT	2373
Db	2101	GCTTTAAATATAAATGGGTCAAAATAGAAATTTGACGATCAGGTGGAAGACTCCAAATAT	2160
Qy	2374	CTAGCTTCTCGATATGATTTCACTTGA CTTAGATCGTGTGGGCATCCAGGCTGCTCTAT	2433
Db	2161	CTAGCTTCTCGATATGATTTCACTTGA CTTAGATCGTGTGGGCATCCAGGCTGCTCTAT	2220
Qy	2434	GGAGGATACCTCTCCCTCATGGCATTAATG CAGAGGTCAGATATCTTTCAGGGTTCATTT	2493
Db	2221	GGAGGATACCTCTCCCTCATGGCATTAATG CAGAGGTCAGATATCTTTCAGGGTTCATTT	2280
Qy	2494	GCTGGGGCCCCAGTCACCTCTGTGGATCTTTCTATGATACAGGATACACGGAAACGTTATATG	2553
Db	2281	GCTGGGGCCCCAGTCACCTCTGTGGATCTTTCTATGATACAGGATACACGGAAACGTTATATG	2340
Qy	2554	GGTCACCTCTGACAGAAATGAACAGGGCTATTACTTTAGGATCTGTGGCCATGCAAGCAGAA	2613
Db	2341	GGTCACCTCTGACAGAAATGAACAGGGCTATTACTTTAGGATCTGTGGCCATGCAAGCAGAA	2400
Qy	2614	AAGTTCCCTCTGAAACAAATCGTTTACTGCTCTTACATGCTTTCTCTGGATGAGAATGTC	2673
Db	2401	NN	2460
Qy	2674	CATTTTGCACATACAGGATATTTACTGAGTTTCTTTAGTGAGGCTGGAAGCCATATGAT	2733
Db	2461	CATTTTGCACATACAGGATATTTACTGAGTTTCTTTAGTGAGGCTGGAAGCCATATGAT	2520
Qy	2734	TTACAGATCTATCTCAGAGAGACACAGCATAGAGTTCTCTGAATCGGAGAACATTTAT	2793
Db	2521	TTACAGATCTATCTCAGAGAGACACAGCATAGAGTTCTCTGAATCGGAGAACATTTAT	2580
Qy	2794	GAATGCACTTTTGGACTACCTTCAGAAACCTTGGATCAGCTATTTGCTGCTTAAAA	2853
Db	2581	GAATGCACTTTTGGACTACCTTCAGAAACCTTGGATCAGCTATTTGCTGCTTAAAA	2640
Qy	2854	GTGATATAA 2862	
Db	2641	GTGATATAA 2649	
RESULT 3			
AK016546		3143 bp mRNA linear	HTC 03-APR-2004
LOCUS		Mus musculus adult male testis cDNA, RIKEN full-length enriched	
DEFINITION		library, clone:4932434F09 product:DIPEPTIDYL PEPTIDASE 8 homolog	
		[Homo sapiens], full insert sequence.	
ACCESSION		AK016546.1 GI:12855334	
VERSION		AK016546	
KEYWORDS		HTC; CAP trapper.	
SOURCE		Mus musculus (house mouse)	
ORGANISM		Mus musculus	
REFERENCE			
AUTHORS		Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,	
JOURNAL		Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayaehizaki, Y.	
PUBLISHED		Normalization and subtraction of cap-trapper-selected cDNAs to	
REFERENCE		prepare full-length cDNA libraries for rapid discovery of new genes	
AUTHORS		Genome Res. 10 (10), 1617-1630 (2000)	
		11042159	
TITLE			
JOURNAL		Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,	
PUBLISHED		Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M.,	
REFERENCE			
AUTHORS			

Qy	918	TGTGCAATAGCTAGCCAAATGGGAAGATGCCAGATCAGCTGGAGTCGCTACCTT	977
Db	1061	CGTACACATAGCTAGCCAAATGGGAAGATGCCAGATCAGCTGGAGTCGCTACCTT	1120
Qy	978	TGTTCTCCAAGAAGAAATTTGATAGATATCTGGCTATTTGGTGTGTCCTCAAAAGCTGAAC	1037
Db	1121	TGTTCTCCAAGAAGAAATTTGATAGATATCTGGCTATTTGGTGTGTCCTCAAAAGCTGAAC	1180
Qy	1038	AACTCCCAAGTGGTGAATAATCTTAGAAATCTATATGAAGAAATATGATGAATCTGAGGT	1097
Db	1181	AACTCCCAAGTGGTGAATAATCTTAGAAATCTATATGAAGAAATATGATGAATCTGAGGT	1240
Qy	1098	GGAAATATTCATGTTATCATCCCTATGTTGGAAACCAAGGAGGCGAGATTCATTCGGTTA	1157
Db	1241	GGAAATATTCATGTTATCATCCCTATGTTGGAAACCAAGGAGGCGAGATTCATTCGGTTA	1300
Qy	1158	TCCTAAACAGGTACAGCAATCTTAAGTCACTTTTAAGATGTCAGAAATATGATGAAT	1217
Db	1301	TCCTAAACAGGTACAGCAATCTTAAGTCACTTTTAAGATGTCAGAAATATGATGAAT	1360
Qy	1218	TCCTGAAGAGATCATAGATGTCATAGATAGGAATTAATCAACCTTTTCAGATCT	1277
Db	1361	TCCTGAAGAGATCATAGATGTCATAGATAGGAATTAATCAACCTTTTCAGATCT	1420
Qy	1278	ATTGAAGAGTTGAATATATGCGAGAGCTGGATGAGCTCCTGAGGGAATATGCTTG	1337
Db	1421	ATTGAAGAGTTGAATATATGCGAGAGCTGGATGAGCTCCTGAGGGAATATGCTTG	1480
Qy	1338	GTCCATCTCTAGATGCTCCAGACTCGCTACAGATAGTGTGATCTCAACCTGAAT	1397
Db	1481	GTCCATCTCTAGATGCTCCAGACTCGCTACAGATAGTGTGATCTCAACCTGAAT	1540
Qy	1398	ATTATCCCAAGTGAAGATGATGTTATGGAAGGCGAGAGCTCATTTAGTTCAGTCTGCA	1457
Db	1541	ATTATCCCAAGTGAAGATGATGTTATGGAAGGCGAGAGCTCATTTAGTTCAGTCTGCA	1600
Qy	1458	TTCTGTGAGCGGCACATTAATCTATGAAGAAACAAAGACATCTGGATAAATCATCA	1517
Db	1601	TTCTGTGAGCGGCACATTAATCTATGAAGAAACAAAGACATCTGGATAAATCATCA	1660
Qy	1518	CATCTTCAATGTTTTCCCAAGTCAAGAGAGAAATTTAGTATTTTGGCTCTCA	1577
Db	1661	CATCTTCAATGTTTTCCCAAGTCAAGAGAGAAATTTAGTATTTTGGCTCTCA	1720
Qy	1578	ATGCAAAACAGGTTTCCTCATTTATATCAAAATTCATCTATTTTAAAGGAAGCAATA	1637
Db	1721	ATGCAAAACAGGTTTCCTCATTTATATCAAAATTCATCTATTTTAAAGGAAGCAATA	1780
Qy	1638	TAAACGATCCAGTGGGCTGCTGCTCCAAAGTATTTCAAGTGTCTTATCAAGAGCA	1697
Db	1781	TAAACGATCCAGTGGGCTGCTGCTCCAAAGTATTTCAAGTGTCTTATCAAGAGCA	1840
Qy	1698	GATAGCAATTTACAGTGGTGAATGGGAAGTTCTTGGCCGCGATGGATCTAATATCCAAT	1757
Db	1841	GATAGCAATTTACAGTGGTGAATGGGAAGTTCTTGGCCGCGATGGATCTAATATCCAAT	1900
Qy	1758	TGATGAAGTCAGAGGCTGGTATATTTTGAAGGCAACCAAGACTCCCTTTTAGAGCATCA	1817
Db	1901	TGATGAAGTCAGAGGCTGGTATATTTTGAAGGCAACCAAGACTCCCTTTTAGAGCATCA	1960
Qy	1818	CTGTACGTGAGTCAAGTGAATATCTGGAGAGTGACAGGCTGACTGACCGTGGCTA	1877
Db	1961	CTGTACGTGAGTCAAGTGAATATCTGGAGAGTGAGTGGTGGTGGTGGTGGTGGTGGT	2020
Qy	1878	CTCACATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1937
Db	2021	CTCACATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2080
Qy	1938	GGAGATCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1997
Db	2081	GGAGATCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2140

RESULT 4

AK029788

LOCUS

DEFINITION

Mus musculus adult male testis cDNA, RIKEN full-length enriched
5517 bp mRNA linear HTC 03-APR-2004

Qy	1998	CAAAACAAAGGAATTTGGGCCACATTTTCGATTCAGAGGTCCTCTTCTGACTATAC	2057
Db	2141	TAACAAAGGAATTTGGGCCACATTTTCGATTCAGAGGTCCTCTTCTGACTATAC	2200
Qy	2058	TCCTCAGAAATTTCTCTTTTGAAGTACTACTGGATTTACATTTGATGGGATGCTCTA	2117
Db	2201	CCCTCAGAAATTTCTCTTTTGAAGTACTACTGGATTTACATTTGATGGGATGCTCTA	2260
Qy	2118	CAAGCCTCATGATCTACAGCCTGGAAATAATCCTACTGCTGTCTATATATGTTGG	2177
Db	2261	TAAGCCTCATGATCTACAGCCTGGAAATAATCCTACTGCTGTCTATATATGTTGG	2320
Qy	2178	TCCTCAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2237
Db	2321	TCCTCAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2380
Qy	2238	CCTAGCCTCTCTAGGTTATGTTGTTAGTATAGACAAACAGGAGGATCTCTGTCACCGAGG	2297
Db	2381	CCTAGCCTCTCTAGGTTATGTTGTTAGTATAGACAAACAGGAGGATCTCTGTCACCGAGG	2440
Qy	2298	GCTTAAATTTGAAGGCGCTTTTAAATATATAAATGGGTCAAAATAGAAATTTGACGATCAAGT	2357
Db	2441	ACTTAAATTTGAAGGCGCTTTTAAATATATAAATGGGTCAAAATAGAAATTTGACGATCAAGT	2500
Qy	2358	GGAGGACTCCCAATATCTAGCTCTCTGATATGATTTTCAATTCAGTTCGTGTCGTCAT	2417
Db	2501	GGAGGACTCCCAATATCTAGCTCTCTGATATGATTTTCAATTCAGTTCGTGTCGTCAT	2560
Qy	2418	CAAGCCTGCTCTATGAGGATACCTCTCCCTGATGTCATTTGATTTGATGATGATGATGAT	2477
Db	2561	CAAGCCTGCTCTATGAGGATACCTCTCCCTGATGTCATTTGATTTGATGATGATGATGAT	2620
Qy	2478	CTTACGGTGTCTATTTGCGGCCCAAGTCACTCTGTCGATCTTCTTATGATACAGGATA	2537
Db	2621	CTTACGGTGTCTATTTGCGGCCCAAGTCACTCTGTCGATCTTCTTATGATACAGGATA	2680
Qy	2538	CACGGAACGTTATATGCGGTACCTGACAGAAATGAACAGGCTATTTACTTTAGGATCTGT	2597
Db	2681	CACGGAACGTTATATGCGGTACCTGACAGAAATGAACAGGCTATTTACTTTAGGATCTGT	2740
Qy	2598	GGCCATGCAAGCAGAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGGTTT	2657
Db	2741	GGCCATGCAAGCAGAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGGTTT	2800
Qy	2658	CCTGGATGAGAATGTCTCATTTTGCACATACCATGATATTTACTGAGTCTTTTGTAGGAGGC	2717
Db	2801	CCTGGATGAGAATGTCTCATTTTGCACATACCATGATATTTACTGAGTCTTTTGTAGGAGGC	2860
Qy	2718	TGGAAGCCATATGATTTACAGATCTATCTTCAGGAGAGACACAGCATTAAGATTCCTGA	2777
Db	2861	TGGAAGCCATATGATTTACAGATCTATCTTCAGGAGAGACACAGCATTAAGATTCCTGA	2920
Qy	2778	ATCGGAGACATATGAACTGCATCTTTTGCACCTTTCAGGAGAAACCTTTGGATCACG	2837
Db	2921	ATCGGAGACATATGAACTGCATCTTTTGCACCTTTCAGGAGAAACCTTTGGATCACG	2980
Qy	2838	TATTGCTGCTCTAAAGTGAATATTTTGAAGTGTGATGATGATGATGATGATGATGATGATGAT	2897
Db	2981	TATTGCTGCTCTAAAGTGAATATTTTGAAGTGTGATGATGATGATGATGATGATGATGATGAT	3036
Qy	2898	CTATTTAAACAAATGAGGAGTTTAAATCAACAGAAACACAGAAATTTGATCATCATTTT	2957
Db	3037	---TCTAACAAACAAAGAGGTTAGCCAAATGGAGAGCAGAGTGTGATGATGATGATGATGATGAT	3093
Qy	2958	GATACCTGCCATGATACATCTACTCTGAAAAATAAATGTTGGTGCATGC	3006
Db	3094	TGTACCTGCCATGATACATCTACTCTGAAAAATAAATGTTGGTGCATGC	3142

library, clone:4930560C15 product:DIPEPTIDYL PEPTIDASE 8 homolog
[Homo sapiens], full insert sequence.

ACCESSION AK029788
VERSION AK029788.1 GI:26081520
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.

1
Carninci,P. and Hayashizaki,Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
10349636

2
Carninci,P. and Hayashizaki,Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
10349636

3
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
11042159

4
Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
Yamamoto,R., Inoue,K., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
Fujiwaka,S., Motome,K., Togawa,Y., Izawa,M., Ohara,E., Watabiki,M.,
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
11076861

5
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)

6
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)

7
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
Hayashida,K., Hayatsu,N., Hiramoto,K., Hirooka,T., Hirozane,T.,
Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
Kato,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Munata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
Muramatsu,M. and Hayashizaki,Y.
Direct Submission

8
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp,
URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)

9
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.jp/

FEATURES
source
URL:http://fantom.gsc.riken.jp/.
Location/Qualifiers
1..5517
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM,DB:4930560C15"
/clone="4930560C15"
/sex="male"
/tissue_type="testis"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
2221..4837
misc_feature
[note="DIPEPTIDYL PEPTIDASE 8 homolog (Homo sapiens)
(SPTR|Q9HBM5, evidence: FASTA, 95.7%ID, 100%length,
match=2649)
putative"]

ORIGIN
Query Match 69.2%; Score 2158.6; DB 4; Length 5517;
Best Local Similarity 85.9%; Pred. No. 0;
Matches 2496; Conservative 0; Mismatches 334; Indels 77; Gaps 6;
QY 201 AGGAAATGCAACATGCGCAGCAATGGAACAGAACAGCTGGGTGTGAGATATTGA 260
DB 2235 AGGAAGATGCAACATGCGCAGCAATGGAACAGAACAGCTGGGTGTGAGATATTGA 2294
QY 261 AACTCGGAGCTGTGAGGAG---NATATTGAATCAGAGTCCGCTAAATTGGAGCCCTT 317
DB 2295 AACTGCGAGGTGTGAGGAGGGAATGGTGAATCAGAGGACCGACCACTGGAGCCCTT 2354
QY 318 TTATGTTGAGCGGTATTCTCGAGTCAGCTTAAAGAGTGTTCGCGATACCAGAAAATA 377
DB 2355 TTATGTCGACGGTATTCTTGGAGTCAGCTGAAAAGCTGCTTGTGATACCAGAAAATA 2414
QY 378 TCATGGCTACATGATGGCTAAGGCACCATCATGATTTTCTGTTTGTGAAGAGGAATGATCC 437
DB 2415 CCACGGCTACATGATGGCTAAGGCACCATCATGATTTTCTGTTTGTGAAGAGGACGATCC 2474
QY 438 AGATGGACCTCATTCAGACAGAAATCTATTACCTTGCCATGTCTGTGTGAGAACAGAGAAA 497
DB 2475 AGATGGCCCTCACTCAGACAGGGTCTATTACCTTGCCATGTCTGTGTGAGAACAGAGAAA 2534
QY 498 TACACTGTTTATTCTGAAATTTCCAAAATCTCAATATAGACAGCAGCTTTAATGCTCTC 557
DB 2535 TACACTATTTTATTCTGAAATCTTAAATCCATCAACAGAGCAGCAGTCTTAATGCTCTTC 2594
QY 558 TTGGAAGCCCTCTTTTGGATCTTTTTCAGCAACACTGGCATATGGAATGTATCTTCGAGA 617
DB 2595 CTGGAAGCCCTCTTTTGGATCTTTTTCAGCGCAGCAGCTAGACTATGGGATGTATCTTCGAGA 2654
QY 618 AGAAGAACTATTAAAGAGAAGAAAACGCAATTTGGACAGTCGGAATTTCTCTTACGATTA 677
DB 2655 GGAAGAGCTACTCAGAGAAGAAAGCGCAATTTGAAACCGTGGGATTCGCGAGCTTATGATTA 2714
QY 678 TCACCAAGGAAGTGGAAACATTTCTGTTTCAAGCCGCTAGTGGAAATTTATTCACGTAAGA 737
DB 2715 CCACCCAGGAAGCGGAACATTTCTGTTTCAAGCTGGCAGCGCATTTACCATATCAAGA 2774
QY 738 TGGAGGCCCAAGAGATTTACGCAACACCTTTAAAGGCCCAATCTAGTGGAAACTAGTTG 797
DB 2775 TGGAGGACCACATGGATTTTACACACAGACTTTTGGGCCCAATTTAGTGGAAACTAGTTG 2834
QY 798 TCCCAACATACGGAATGATCCAAAATTTATGCCCGCTGATTCAGACTGGATGCTTTTAT 857
DB 2835 TCCCAATATACGAATGATCCAAAATTTATGCCCTGCTGATTCAGACTGGATGCTTTTAT 2894
QY 858 ACATAGCAACGATTTTGGATATCTTAACATCTTAACAGAGAGAGAAAGAGACTCATT 917
DB 2895 TCACAGCAATGATTTTGGATATCAAAACCTCGTAACAGGGAAGGAGGAGTACATA 2954
QY 918 TGTCCCAATGAGCTAGCCCAACATGGAAGAGATGCCAGATCAGCTGGAGTCGCTACCTT 977


```

RESULT 5
LOCUS AY411617
DEFINITION AY411617 2634 bp DNA linear GSS 16-DEC-2003
MUS musculus DPP8 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY411617
VERSION AY411617.1 GI:39767585
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Murioidea; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2634)
AUTHORS Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 2634)
AUTHORS Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source location/Qualifiers
1..2634
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>2634
/gene="DPP8"
/locus_tag="HCM4283"
gene
ORIGIN
Query Match 63.6%; Score 1985; DB 10; Length 2634;
Best Local Similarity 82.4%; Pred. No. 0;
Matches 2185; Conservative 0; Mismatches 446; Indels 21; Gaps 2;
QY 214 ATGGCAGCAGCAATGGAACACAGACAGCTGGGTGTTCAGATATTTGAACTGCGGACTGT 273
DB 1 ATGGCAGCAGCAATGGAACACAGACAGCTGGGTGTTCAGATATTTGAACTGCGGACTGT 60
QY 274 GAGGAG---AATATTGAATCAAGATCGGCTAAATGGAGCCTTTTATGTGAGCGG 330
DB 61 GAGGAGGGAATGGTGAATCAAGGACCGACCCAACTGGAGCCTTTTATGTGCGAACGG 120
QY 331 TATTCCTGAGTCAGCTTAAAGCTGCTCCCATACCGAGAAATATCATGCTGCTACATG 390
DB 121 TATTCCTGAGTCAGCTGAAAGCTGCTGCTGATACCGAGAAATACCGGCTACATG 180
QY 391 ATGGCTAAGGCACCATGATTTTCATGTTTGTGAAGAGGATGATCCAGATGACCTCAT 450
DB 181 ATGGCTAAGGCACCATGATTTTCATGTTTGTGAAGAGGATGATCCAGATGACCTCAT 240
QY 451 TCAGACAGCAATCTATTACCTTGCCATGTCTGGTGAGAACACAGAGAAATACATGTTTAT 510
DB 241 TCAGACAGCGGTCTATTACCTTGCCATGTCTGGTGAGAACACAGAGAAATACATGTTTAT 300
QY 511 TCTGAATTTCCCAAACTATCAATAGACAGAGCTTTTATGCTCTCTTGGAGGCTCTT 570
DB 301 TCTGAATTTCCCAAACTATCAATAGACAGAGCTTTTATGCTCTCTTGGAGGCTCTT 360
QY 571 TTGGATCTTTTTCAGGCAACACTGGACTATGGAATGTATTCTCGAGAGAGAACTATTATTA 630

```

```

Db 361 TTGGATCTTTTTCAGGCGACCTAGACTATGGATGTATTCTCGAGAGAGAGAGCTACTC 420
QY 631 AGAGAAGAAACGCAATTCGAAACAGTCGGAATTTGCTTCTTACGATTTATCACAAGGAAGT 690
Db 421 AGAGAAGAAACGCAATTCGAAACGTCGGAATTTGCTTCTTACGATTTATCACAAGGAAGT 480
QY 691 GGAACATTTCTGTTTCAAGCGGTAGTGAATTTATCAGTAAAGAGATGGAGGGCCACAA 750
Db 481 GGAACATTTCTGTTTCAAGCGGTAGTGAATTTATCAGTAAAGAGATGGAGGGCCACAT 540
QY 751 GGAATTTACGCAACCAACTTTTAAGGCCCAATCTAGTGGAAACTAGTGTGTCACCAATACGG 810
Db 541 GGAATTTACCAACAGCGCTTTGCGGCCCAATTTAGTGGAAACTAGTGTGTCACCAATACGA 600
QY 811 ATGGATCCCAAAATTTATGCCCCGCTGATCCAGACTGGATGCTTTTATCATAGCAACGAT 870
Db 601 ATGGATCCCAAAATTTATGCCCCGCTGATCCAGACTGGATGCTTTTATCATAGCAACGAT 660
QY 871 ATTTGGATATCTAACATCTGTAACCAAGAGAAAGAGAGACTCCTATTGTGACAAATGAG 930
Db 661 ATTTGGATATCAAACTCTGTAACCAAGAGAGAGAGAGACTCCTATTGTGACAAATGAG 720
QY 931 CTAGCCCAACATGGAAGAGATGCCAGTCCAGTGGAGTGGCTTACCTTTGTTCTCCAGAA 990
Db 721 CTAGCCCAACATGGAAGAGATGCCAGTGGAGTGGCTTACCTTTGTTCTCCAGAA 780
QY 991 GAATTTGATAGATATTCTGGCTATTGGTGTCTGTCGAAAGCTGAAACAACTCCAGTGGT 1050
Db 781 GAATTTGACAGATCTCTGGCTACTGGTGTCTGTCGAAAGCTGAAACAACTCCAGTGGT 840
QY 1051 GGTAAATTTCTTGAATTTCTATATGAAGAAATGATGATCTGAGGTGGAATATTATTCAT 1110
Db 841 GGTAAATTTCTTGAATTTCTATATGAAGAAATGATGATCTGAGGTGGAATATTATTCAT 900
QY 1111 GTTACATCCCTATGTTGGAACCAAGAGGCGAGATTTCAATCCGTTATCTTAAACACAGT 1170
Db 901 GTTACGTCCTCCCAATGTTGGAACCAAGAGGCGAGATTTCTTTCGTTATCTCCAAACAGG 960
QY 1171 ACAGCAATCTTAAAGTCACTTTTAAGATGTCAGAAATAATGATGATGCTGGAAGGAAG 1230
Db 961 ACAGCAATCTTAAAGTCACTTTTAAGATGTCAGAAATAATGATGATGCTGGAAGGAAG 1020
QY 1231 ATCATAGATGTCATAGATAAGGAATAATTTCAACCTTTTGAGATTTCTATTTGAAAGGAT 1290
Db 1021 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1080
QY 1291 GAATATATTGCCAGAGCTGGATGGACTCTGAGGGAATAATATGCTTGGTCCATCTCTACTA 1350
Db 1081 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1140
QY 1351 GATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGGAATTTATTTATCCAGTA 1410
Db 1141 GACCGTTCACAGACTCACCTGCGAGATGTTCTGATCTCCCTCGAGTATTTATCCAGTA 1200
QY 1411 GAAGATGATGTTTATGAAAGGCGAGACTCATTTAGTCACTGCTGATCTTGTGACGCCA 1470
Db 1201 GAAGATGATGTTTATGAAAGGCGAGACTCATTTAGTCACTGCTGATCTTGTGACGCCA 1260
QY 1471 CTAAATATTCTATGAAGAAACAAACAGACATCTCGATAAATAATTCATGACATCTTTTCA 1530
Db 1261 CTGATCATCTATGAAGAAACAAACAGACATCTCGATAAATAATTCATGACATCTTTTCA 1320
QY 1531 TTTCCCAAGTCCAGAGGGAATTTAGTGTATTTTGTGCTGCTGAATGCAAAACAGT 1590
Db 1321 TTTCTCAAACTCATGAAGATGAAATTTAGTGTATTTTGTGCTGCTGAATGCAAAACAGT 1380
QY 1591 TTTCTCAAACTCATGAAGATGAAATTTAGTGTATTTTGTGCTGCTGAATGCAAAACAGT 1650
Db 1381 TTTCTCAAACTCATGAAGATGAAATTTAGTGTATTTTGTGCTGCTGAATGCAAAACAGT 1440
QY 1651 GGTGGGCTGCTGCTCCAAAGTGAATTTCAAGTGTCTCTATCAAGAGGAGATAGCAATTACC 1710

```


Db 249 CCGGTAGTGGAAATTTATCACGTTAAAGATGGAGGCCACAGGATTTACGCAACACCTTT 308
Qy 770 TAAGGCCAACTTAGTGGAAACTAGTTGTCCTCAACATACGGATGGATCCAAAATTTAGCC 829
Db 309 TAAGGCCAACTTAGTGGAAACTAGTTGTCCTCAACATACGGATGGATCCAAAATTTAGCC 368
Qy 830 CGGCTGATCCAGACTGGATGCTTTTATACATAGCAACGATATTTGGATATCTAAACATCG 889
Db 369 CTGCTGATCCAGACTGGATGCTTTTATACATAGCAACGATATTTGGATATCTAAACATCG 428
Qy 890 TAACACAGAGAAGAGGAGACTCATTATGTGCACAATGAGCTAGCCCAACATCGGAAGAAG 949
Db 429 TAACACAGAGAAGAGGAGACTCATTATGTGCACAATGAGCTAGCCCAACATCGGAAGAAG 488
Qy 950 ATGCCAGATCAGCTGGAGTCGCTACCTTTGTCTCCAAAGAAGAAATTTGATAGATATTCG 1009
Db 489 ATGCCAGATCAGCTGGAGTCGCTACCTTTGTCTCCAAAGAAGAAATTTGATAGATATTCG 548
Qy 1010 GCTATTTGGTGTCTCCAAAGCTGAACAACTCCACAGTGGTGTGTTAAATTTCTTAGAATTC 1069
Db 549 GCTATTTGGTGTCTCCAAAGCTGAACAACTCCACAGTGGTGTGTTAAATTTCTTAGAATTC 608
Qy 1070 TATATGAAGAAATGATGAATCTGAGGTGGAAATTTATTCATGTTTACATCCCTATCTGTG 1129
Db 609 TATATGAAGAAATGATGAATCTGAGGTGGAAATTTATTCATGTTTACATCCCTATCTGTG 668
Qy 1130 AAACAAGGAGGCGAGATTCATTCCGTTATCTTAAACACAGGTACAGCAAAATCTTAAAGTCA 1189
Db 669 AAACAAGGAGGCGAGATTCATTCCGTTATCTTAAACACAGGTACAGCAAAATCTTAAAGTCA 728
Qy 1190 CTTTTAAGATGTCAGAAATATGATGATGCTGTAAGGAAGATCATAGATGTCATAGATA 1249
Db 729 CTTTTAAGATGTCAGAAATATGATGATGCTGTAAGGAAGATCATAGATGTCATAGATA 788
Qy 1250 AGGAATAATCAACCTTTTGAGATCTTATTTGAGGAGTTGAATATATTTGCCAGAGCTG 1309
Db 789 AGGAATAATCAACCTTTTGAGATCTTATTTGAGGAGTTGAATATATTTGCCAGAGCTG 848
Qy 1310 GATGACCTCTGAGGGAATAATGCTGGTCCATCTTACTAGATCGCTCCAGACTCGCC 1369
Db 849 GATGACCTCTGAGGGAATAATGCTGGTCCATCTTACTAGATCGCTCCAGACTCGCC 908
Qy 1370 TACAGATAGTGTGATCTCAGCTGAATTTATTTATCCAGTAGAAGATGATGTTATCGAAA 1429
Db 909 TACAGATAGTGTGATCTCAGCTGAATTTATTTATCCAGTAGAAGATGATGTTATCGAAA 968
Qy 1430 GGCAGAGACTCAATGAGTCAGTGCCTGATCTCTGTGACGCCACTAAATTAATCTATGAAGAAA 1489
Db 969 GGCAGAGACTCAATGAGTCAGTGCCTGATCTCTGTGACGCCACTAAATTAATCTATGAAGAAA 1028
Qy 1490 CAACAGACATCTGGATAAATATCCATGACATCTTTCAATGTTTTTCCCAAGTCAGGAAG 1549
Db 1029 CAACAGACATCTGGATAAATATCCATGACATCTTTCAATGTTTTTCCCAAGTCAGGAAG 1088
Qy 1550 AGGAAATTTAGTGTATTTTTTGGCTCTGAAATGCAAAACAGGTTTCCGTCATTTATACAAA 1609
Db 1089 AGGAAATTTAGTGTATTTTTTGGCTCTGAAATGCAAAACAGGTTTCCGTCATTTATACAAA 1148
Qy 1610 TTACATCTATTTTAAAGGAAGCAAAATATAAACATCCAGTGGTGGCTGCTGCTCCAA 1669
Db 1149 TTACATCTATTTTAAAGGAAGCAAAATATAAACATCCAGTGGTGGCTGCTGCTCCAA 1208
Qy 1670 GTGATTTCAAGTGCTCTATCAAGAGGAGATAGCAATTTACCAGTGGTGAATGGGAAGTTC 1729
Db 1209 GTGATTTCAAGTGCTCTATCAAGAGGAGATAGCAATTTACCAGTGGTGAATGGGAAGTTC 1268
Qy 1730 TTGGCGGCGATCGGATCTAATATCCAAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAG 1789
Db 1269 TTGGCGGCGATCGGATCTAATATCCAAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAG 1328
Qy 1790 GCACCAAGACTCCCTTTTAGAGCATCACCTGTACGTAGTCAGTTACGTTAAATCTCTGGAG 1849
Db 1329 GCACCAAGACTCCCTTTTAGAGCATCACCTGTAGTCAGTTACGTTAAATCTCTGGAG 1388

Qy 1850 AGGTGACAGAGCTGACTGACCGTGGCTACTCAATCTTGTCTGCAATCAGTCAGCACTGTG 1909
Db 1389 AGGTGACAGAGCTGACTGACCGTGGCTACTCAATCTTGTCTGCAATCAGTCAGCACTGTG 1448
Qy 1910 ACTTCTTTTATAAGTAAAGTATAGTAACCAAGAAATCCACACTGCTGTGTCCTTTTACAAGC 1969
Db 1449 ACTTCTTTTATAAGTAAAGTATAGTAACCAAGAAATCCACACTGCTGTGTCCTTTTACAAGC 1508
Qy 1970 TATCAAGTCTGAAAGTGAACCCAACTTGCAAAACAAAGGAATTTTGGGCCACCAATTTGG 2029
Db 1509 TATCAAGTCTGAAAGTGAACCCAACTTGCAAAACAAAGGAATTTTGGGCCACCAATTTGG 1568
Qy 2030 ATTACAGAGTCTCTTCTCTGACTATATCTCTCCAGAAATTTTCTTCTTTGAAAGTACTA 2089
Db 1569 ATTACAGAGTCTCTTCTCTGACTATATCTCTCCAGAAATTTTCTTCTTTGAAAGTACTA 1628
Qy 2090 CTGGATTTACATTTGATGGATGCTCTCAAGAGCTCATGATCTACAGCTCGGAAGAAT 2149
Db 1629 CTGGATTTACATTTGATGGATGCTCTCAAGAGCTCATGATCTACAGCTCGGAAGAAT 1688
Qy 2150 ATCTACTGTGCTGTTTATATATATGTTGTTCTCTCAGTGCAGTTCAGTTCGTTA 2209
Db 1689 ATCTACTGTGCTGTTTATATATGTTGTTCTCTCAGTGCAGTTCAGTTCGTTA 1722
Qy 2210 AAGGAGTCAAGTATTTCCGTTGAATAACCTTAGCCTCTCTAGGTTATGTTGTTAGTGA 2269
Db 1723 AAGGAGTCAAGTATTTCCGTTGAATAACCTTAGCCTCTCTAGGTTATGTTGTTAGTGA 1722
Qy 2270 TAGACAACAGGGGATCCTGTACCCGAGGCTTAAATTTGAAGGCGCTTTAAATATAAAA 2329
Db 1723 TAGACAACAGGGGATCCTGTACCCGAGGCTTAAATTTGAAGGCGCTTTAAATATAAAA 1722
Qy 2330 TGGGTCAAAATAGAAAATGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATAG 2389
Db 1723 TGGGTCAAAATAGAAAATGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATAG 1722
Qy 2390 ATTTCAATTGACTTAGATCGTGTGGGCATCCAGGCTGGTCTTATGAGGATACCTCTCCC 2449
Db 1723 ATTTCAATTGACTTAGATCGTGTGGGCATCCAGGCTGGTCTTATGAGGATACCTCTCCC 1722
Qy 2450 TGATGGCATTAATGACAGAGTCAGATATCTTCAAGGGTTGCTATTTGCTGGGGCCCCAGTCA 2509
Db 1723 TGATGGCATTAATGACAGAGTCAGATATCTTCAAGGGTTGCTATTTGCTGGGGCCCCAGTCA 1748
Qy 2510 CTCTGTGGATCTTCTATGATACAGGATACAGGAAGCTTATATGCTGTCACCTGACCCAGA 2569
Db 1749 CTCTGTGGATCTTCTATGATACAGGATACAGGAAGCTTATATGCTGTCACCTGACCCAGA 1808
Qy 2570 ATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGAAC 2629
Db 1809 ATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGAAC 1868
Qy 2630 CAAATCGTTTACTGCTCTTACATGTTTCTGATGAGAAATGTCATTTTGCACATACCA 2689
Db 1869 CAAATCGTTTACTGCTCTTACATGTTTCTGATGAGAAATGTCATTTTGCACATACCA 1928
Qy 2690 GTATATTTACTGAGTTTTTTAGTGGGCTGGGAAGCCATATGATTTTACAGATCTATCCTC 2749
Db 1929 GTATATTTACTGAGTTTTTTAGTGGGCTGGGAAGCCATATGATTTTACAGATCTATCCTC 1976
Qy 2750 AGGAGAGACACAGCATAAAGATTTCTGAATCGGAGAAACATTTGAACATGCTCTTTTGC 2809
Db 1977 AGGAGAGACACAGCATAAAGATTTCTGAATCGGAGAAACATTTGAACATGCTCTTTTGC 2036
Qy 2810 ACTACCTTCAAGAAAACCTTTGGATCAGGATTTGCTGCTCTTAAAGTATATAATTTTGC 2869
Db 2037 ACTACCTTCAAGAAAACCTTTGGATCAGGATTTGCTGCTCTTAAAGTATATAATTTTGC 2096
Qy 2870 CTGTGTAGAACTCTCTGGTATACACTGGCTATTTTAAACCAATGAGGAGGTTTATCAACA 2929
Db 2097 CTGTGTAGAACTCTCTGGTATACACTGGCTATTTTAAACCAATGAGGAGGTTTATCAACA 2156

Tue Apr 18 08:18:38 2006

QY	2930	GAACAACAGAAATGATCATCAATTTTGGATGCTGCTACAGGCTCATGATCTACA	2989	2084	GTACTACTGGAATTTACATTTGATGGGATGCTCTACAGGCTCATGATCTACA	2140
DB	2157	GAACAACAGAAATGATCATCAATTTTGGATGCTGCTACAGGCTCATGATCTACA	2216	241	GTACTACTGGAATTTACATTTGATGGGATGCTCTACAGGCTCATGATCTACA	300
QY	2990	TAAATGTTGGTCCATGACAGGGCTCAGGTTTGGTGTAGTATCTAATACCTTAACCCCA	3049	2141	GAAGAATAATCTCACTGT- GCTGTTTATATATGTTGGTCTCAGGTCAGTTGGTGAAT	2199
DB	2217	TAAATGTTGGTCCATGACAGGGCTCAGGTTTGGTGTAGTATCTAATACCTTAACCCCA	2276	301	GAAGAATAATCTCACTGTGNGGCTGTTTATATATGAGGGCCCTAAGTGAGTTGGTGAAT	360
QY	3050	CATGCTCAAAATCAAA 3065		2200	AATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATAACCTTACCTCTCTAGGTTATGTG	2259
DB	2277	CATGCTCAAAATCAAA 2292		361	AATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATAACCTTACCTCTCTAGGTTATGTG	420
RESULT 7				2260	GTGTTAGTGTAGACAAACAGGGGATCCTGTCAACGAGGGCTTAAATTTGAAGCGGCTTT	2319
LOCUS	AF175225	1292 bp mRNA linear HTC 01-AUG-2003		421	GTGTTAGTGTAGACAAACAGGGGATCCTGTCAACGAGGGCTTAAATTTGAAGCGGCTTT	480
DEFINITION	Homo sapiens tissue-type aorta MSTP135 mRNA, complete cds.			2320	AAATATAAAATGGGTCAAAATAGAAATGACGATCAGGTGGAAGGACTCCAATATCTAGCT	2379
ACCESSION	AF175225			481	AAATTTAAATGGGTCAAAATAGAAATGACGATCAGGTGGAAGGACTCCAATATCTAGCT	540
VERSION	AF175225.1	GI:33338055		2380	TCTCGATATGATTTTCAATTTGACTTAGATCGTGTGGGCATCCACGGCTGGTCTCTATGGAGGA	2439
KEYWORDS	HTC.			541	TCTCGATATGATTTTCAATTTGACTTAGATCGTGTGGGCATCCACGGCTGGTCTCTATGGAGGA	600
SOURCE	Homo sapiens (human)			2440	TACCTCTCCCTGATGCGCATTAATGACAGAGTCAAGTATCTTTCAGGGTGGTCTATTTGCTGGG	2499
ORGANISM	Homo sapiens			601	TACCTCTCCCTGATGCGCATTAATGACAGAGTCAAGTATCTTTCAGGGTGGTCTATTTGCTGGG	660
REFERENCE	1 (bases 1 to 1292)			2500	GCCCCAGTCACTCTCTGGATCTTCTATGATACAGGATACACGGAACGTTATATGGGTGAC	2559
AUTHORS	Zhao, B., Xu, H. S., Tong, Y. K., Sheng, H., Qin, B. M., Liu, Y. Q., Liu, B., Wang, X. Y., Zhang, Q., Song, L., Gao, Y., Zhang, C. L., Ye, J., Ji, X. J., Liu, B. H., Lu, H., Chen, J. Z., Cai, M. Q., Zheng, W. Y., Teng, C. Y., Liu, Q., Yu, L. T., Lin, J., Gong, Q., Zhang, A. M., Gao, R. L. and Hui, R. T.			661	GCCCCAGTCACTCTCTGGATCTTCTATGATACAGGATACACGGAACGTTATATGGGTGAC	720
TITLE	Direct Submission			2560	CTGACACAGATGAACAGGGCTTATCTTAGATCGTGTGGGCATCCACGGCTGGTCTCTATGGGTG	2619
JOURNAL	Submitted (04-AUG-1999) Molecular Medicine Center for Cardiovascular Disease, Cardiovascular Institute, CAMS & PUMC, 167, Bei Li Shi Lu, Beijing 100037, P.R. China			721	CCTGACACAGATGAACAGGGCTTATCTTAGATCGTGTGGGCATCCACGGCTGGTCTCTATGGGTG	780
FEATURES	Location/Qualifiers			2620	CCCTCTGAACAAATCGTTTACTGCTCTTACATGTTTCTGCTGATGAGATGTCATTTT	2679
source	1. .1292			781	CCCTCTGAACAAATCGTTTACTGCTCTTACATGTTTCTTGGATGAGATGTCATTTT	840
CDS	/organism="Homo sapiens"			2680	GCACATACCATGATATTTACTGATTTTCTGATGAGGCTGGAAGCCATGATTTTACAG	2739
	/mol_type="mRNA"			841	GCACATACCATGATATTTACTGATTTTCTGATGAGGCTGGAAGCCATGATTTTACAG	900
	/db_xref="taxon:9606"			2740	ATCTATCTCAGGAGAGACACAGCATTAAGAGTTCTTGAATCGGGAGAACATTTGAACTG	2799
	/tissue_type="aorta"			901	ATCTATCTCAGGAGAGACACAGCATTAAGAGTTCTTGAATCGGGAGAACATTTGAACTG	960
	142. .1023			2800	CATCTTTTGACCTACCTTCAAGAAAACTTGGATCACGTTATTTGCTCTTAAAGTGATA	2859
	/product="MSTP135"			961	CATCTTTTGACCTACCTTCAAGAAAACTTGGATCACGTTATTTGCTCTTAAAGTGATA	1020
	/protein_id="AAQ13650.1"			2860	TAAATTTGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTTTAAACCAATGAGGAGT	2919
	/translation="MTQAKQRFPPWQVLFLLILQKSLKLVLLDLHMCSS			1021	TAAATTTGACCTGTGTAGAACTCTCTGGTATACACTGGCTATTTTAAACCAATGAGGAGT	1080
	TSFHSSTGLERNILKLFYGGPKVQLVNNRFGVFRNLTPFLGYVXGIDNRGS			2920	TAAATCAACAGAAACACAGAAATTTGATCATCATTTTGCATCTGCGCATGTAACATCTA	2979
	CPRGLKFEAGFKMGQIEIDDOVEGLQYLASRYDFDLDRVGHGWSYGLSLMAL			1081	TAAATCAACAGAAACACAGAAATTTGATCATCATTTTGCATCTGCGCATGTAACATCTA	1140
	MORSDFRVAIAGAPVTLWIFDYTYTERYMGHPDQNEQSYILGSVAMQREKFPSEPN			2980	CTCTGAAAATAAATGTTGGTGCCTCAGGGGTCTACGGTTTGTGTGTAGTAACTTAAATAC	3039
	RLILLHGFUDENVHFAHTSILLFLVRAGKPYDLQIYQERHSIRVPSGEHYELHLL			1141	CTCTGAAAATAAATGTTGGTGCCTCAGGGGTCTACGGTTTGTGTGTAGTAACTTAAATAC	1200
	HVLLQENLGSRIAALKVI"			3040	CTTAAACCCCAATCTCTCAAAATCAAAATGATACATTTCTCTGAGAGACCCAGCAATACCAT	3099
ORIGIN				1201	CTTAAACCCCAATCTCTCAAAATCAAAATGATACATTTCTCTGAGAGACCCAGCAATACCAT	1260
Query Match	39.0%; Score 1217.4; DB 4; Length 1292;			3100	AAGAATTAATAAAAAAAA 3120	
Best Local Similarity	97.9%; Pred. No. 0;			1261	AAGAATTAATAAAAAAAA 1281	
Matches 1254; Conservative	0; Mismatches 23; Indels 4; Gaps 2;					
QY	1844	CTGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACATTTCTGCTGATCAGTCAGC	1903			
DB	1	CTGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACATTTCTGCTGATCAGTCAGC	60			
QY	1904	ACTGTGACTCTTTTAAAGTATAGTAAACAGAGAAATCCACACTGTGTGTCCTTTT	1963			
DB	61	ACTGTGACTCTTTTAAAGTATAGTAAACAGAGAAATCCACACTGTGTGTCCTTTT	120			
QY	1964	ACAAGCTATCAGTCTGAGATGACCCCACTTGCACAAAGAAATTTTGGGCCACCA	2023			
DB	121	ACAAGCTATCAGTCTGAGATGACCCCACTTGCACAAAGAAATTTTGGGCCACCA	180			
QY	2034	TTTGGATTCAGAGGTCCTCTCTGATCTACTCTCCAGAAATTTCTCTTTGAAA	2083			
DB	181	TTTGGATTCAGAGGTCCTCTCTGATCTACTCTCCAGAAATTTCTCTTTGAAA	240			


```
RESULT 8
AF176779
LOCUS      Homo sapiens MSTP141 mRNA, complete cds.          HTC 01-AUG-2003
DEFINITION
ACCESSION  AF176779
VERSION     AF176779.1  GI:33338069
KEYWORDS   HTC.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Homiidae; Homo
REFERENCE   1 (bases 1 to 1265)
AUTHORS     Hui,R.T., Liu,Y.Q., Wang,X.Y., Qin,B.M. and Sheng,H.
TITLE       Homo sapiens normal aorta mRNA MST141
JOURNAL     Unpublished
REFERENCE   2 (bases 1 to 1265)
AUTHORS     Hui,R.T., Liu,Y.Q., Wang,X.Y., Qin,B.M. and Sheng,H.
TITLE       Direct Submission
JOURNAL     Submitted (10-AUG-1999) Molecular Medicine Center for
            Cardiovascular Disease, Cardiovascular Institute, CAMS & PUMC, 167,
            Bei Li Shi Lu, Beijing 100037, P.R. China
FEATURES   Location/Qualifiers
            source          1..1265
                        /organism="Homo sapiens"
                        /mol_type="mRNA"
                        /db_xref="taxon:9606"
                        /tissue_type="aorta"
                        268..741
                        /codon_start=1
                        /product="MSTP141"
                        /protein_id="AAQ13657.1"
                        /db_xref="GI:33338070"
                        /translators="MLV,KPHDLOPKKYPVLVIYGGPOVLVNNRFGVKYFRLNTL
                        ASLGVVVVINDRSGRHLKFEKAPKMGQIEIDQVEGLQYLASRYDFIDLDRVG
                        IHGWSYGYLSLMALMQRSDIFRVAIAGAPVLIWIFDITGTERHMGHPDQEQ"
CDS
ORIGIN
Query Match      38.5%; Score 1200; DB 4; Length 1265;
Best Local Similarity 97.7%; Pred. No. 0;
Matches 1248; Conservative 0; Mismatches 5; Indels 25; Gaps 2;
QY 1844 CTGGAGAGGTGACAAAGG-CTGACTGACCGTGGCTACTCACATTTCTTGCTGCATCAGTCAG 1902
DB 1 CTGGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCACATTTCTTGCTGCATCAGTCAG 60
QY 1903 CACTGTGACTCTTTTAAAGTAAGTATAGTAACAGAGAAGATCCACACTGTGTGCTCTT 1962
DB 61 CACTGTGACTCTTTTAAAGTAAGTATAGTAACAGAGAAGATCCACACTGTGTGCTCTT 120
QY 1963 TACAAGCTATCAAGTCTCGAGATGACCCCACTTGCACAAACAAAGGAATTTTGGGCCACC 2022
DB 121 TACAAGCTATCAAGTCTCGAGATGACCCCACTTGCACAAACAAAGGAATTTTGGGCCACC 180
QY 2023 ATTTTGGATTTCAGCAGGTCTCTTCTGACTATATCTCTCCAGAAAATTTTCTCTTTGAA 2082
DB / 181 ATTTTGGATTTCAGCAGGTCTCTTCTGACTATATCTCTCCAGAAAATTTTCTCTTTGAA 240
QY 2083 AGTACTACTGGATTATTCATTTGATGGGATGCTCTACAAAGCCTCATGATCTACAGCTGGA 2142
DB 241 AGTACTACTGGATTATTCATTTGATGGGATGCTCTACAAAGCCTCATGATCTACAGCTGGA 300
QY 2143 AGAAATATCTCTACTGTGCTGTTTATATATGTTGTTCTCAGGTGCTCAGGTGCTGTTGTTGTT 2202
DB 301 AGAAATATCTCTACTGTGCTGTTTATATATGTTGTTCTCAGGTGCTCAGGTGCTGTTGTTGTT 360
QY 2203 CGTTTAAAGGAGTCAAGTATTTTCGCTTGAATACCTTAGCTCTCTAGGTATGTTGGTT 2262
DB 361 CGTTTAAAGGAGTCAAGTATTTTCGCTTGAATACCTTAGCTCTCTAGGTATGTTGGTT 420
QY 2263 GTAGTGATAGACAAACAGGGGATCTCTGTCCAGCGAGGCTTAAATTTGAAGCGGCTTTAAA 2322
DB 421 GTAGTGATAGACAAACAGGGGATCTCTGTCCAGCGAGGCTTAAATTTGAAGCGGCTTTAAA 480
```

```
QY 2323 TATAAATCGGTCAAAATGAAATTTGACCATCAGGTGGAGGACTCCCAATATCTAGCTTCT 2382
DB 481 TATAAATCGGTCAAAATGAAATTTGACCATCAGGTGGAGGACTCCCAATATCTAGCTTCT 540
QY 2383 CGATATGATTTTCATTTGACTTAGATCGTGTGGGCATCCACGGCTGGTCTCTATGGAGGATAC 2442
DB 541 CGATATGATTTTCATTTGACTTAGATCGTGTGGGCATCCACGGCTGGTCTCTATGGAGGATAC 600
QY 2443 CTCCTCCCTGATGGCATTAATGACAGAGGTGAGATATCTTTCAGGGTGTCTATTTGCTGGGGCC 2502
DB 601 CTCCTCCCTGATGGCATTAATGACAGAGGTGAGATATCTTTCAGGGTGTCTATTTGCTGGGGCC 660
QY 2503 CAGTCACTCTGTGGATCTTCTATGATACAGATACACGGAACTTATATGGGTGACCCCT 2562
DB 661 CAGTCACTCTGTGGATCTTCTATGATACAGATACACGGAACTTATATGGGTGACCCCT 720
QY 2563 GACCAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAGAAAAGTTCCCC 2622
DB 721 GACCAGATGAACAGGGCT-----AAAAAAAAAAGTTCCCC 756
QY 2623 TCTGAACCAAAATCGTTTACTGCTCTTACATGCTTTCTCTGGATGAGAATGTCCATTTGCA 2682
DB 757 TCTGAACCAAAATCGTTTACTGCTCTTACATGCTTTCTCTGGATGAGAATGTCCATTTGCA 816
QY 2683 CATACCAAGTATATTTACTGAGTGTCTTACATGAGGCTGGAAGCCATATGATTTACAGATC 2742
DB 817 CATACCAAGTATATTTACTGAGTGTCTTACATGAGGCTGGAAGCCATATGATTTACAGATC 876
QY 2743 TATCCTCAGAGAGACACAGCATAGAGTTCTCTGAATCGGAGAACATTTATGAATCTGCAT 2802
DB 877 TATCCTCAGAGAGACACAGCATAGAGTTCTCTGAATCGGAGAACATTTATGAATCTGCAT 936
QY 2803 CTTTTGCACCTACCTTCAAGAAAACCTTGGATCACGTTATTTGCTGTCTTAAAGAGTGATATA 2862
DB 937 CTTTTGCACCTACCTTCAAGAAAACCTTGGATCACGTTATTTGCTGTCTTAAAGAGTGATATA 996
QY 2863 TTTTGACCTGTGTAGAACTCTCTGTGTATACATGCTGCTATTAAACCAATGAGGAGTTTA 2922
DB 997 TTTTGACCTGTGTAGAACTCTCTGTGTATACATGCTGCTATTAAACCAATGAGGAGTTTA 1056
QY 2923 ATCAACAGAAAACACAGATTCATCACAATTTTGTATGATCCCTGATGATACATCTACTC 2982
DB 1057 ATCAACAGAAAACACAGATTCATCACAATTTTGTATGATCCCTGATGATGATCACTACTC 1116
QY 2983 CTGAAATAAATGTGTGTCATGCGGGGTCTACCGTTTGTGTAGTAACTTAATACCTT 3042
DB 1117 CTGAAATAAATGTGTGTCATGCGGGGTCTACCGTTTGTGTAGTAACTTAATACCTT 1176
QY 3043 AACCCACATGCTCAAAATCAAAATGATACATATTCCTGAGAGACCCAGCAATACATAAG 3102
DB 1177 AACCCACATGCTCAAAATCAAAATGATACATATTCCTGAGAGACCCAGCAATACATAAG 1236
QY 3103 AATTACTAAAAA 3120
DB 1237 AATTACTAAAAA 1254
```

```
RESULT 9
AF173382
LOCUS      Homo sapiens MSTP097 (MST097) mRNA, complete cds.          HTC 01-AUG-2003
DEFINITION
ACCESSION  AF173382
VERSION     AF173382.1  GI:33338001
KEYWORDS   HTC.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Homiidae; Homo
REFERENCE   1 (bases 1 to 1278)
AUTHORS     Xu,H.S., Zhao,B., Tong,Y.K., Sheng,H., Qin,B.M., Liu,Y.Q., Liu,B.,
            Wang,X.Y., Zhang,Q., Song,L., Gao,Y., Zhang,C.L., Ye,J., Ji,X.J.,
```


TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 Liu, B.H., Lu, H., Chen, J.Z., Cai, M.Q., Zheng, W.Y., Teng, C.Y.,
 Liu, Q., Yu, L.T., Lin, J., Gong, Q., Zhang, A.M., Gao, R.L. and Hiu, R.T.
 Homo sapiens normal aorta mRNA MST097
 2 (bases 1 to 1278)

TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 Xu, H.S., Zhao, B., Tong, Y.K., Sheng, H., Qin, B.M., Liu, Y.Q., Liu, B.,
 Wang, X.Y., Zhang, Q., Song, L., Gao, Y., Zhang, C.L., Ye, J., Ji, X.J.,
 Liu, B.H., Chen, J.Z., Cai, M.Q., Zheng, W.Y., Teng, C.Y.,
 Liu, Q., Yu, L.T., Lin, J., Gong, Q., Zhang, A.M., Gao, R.L. and Hiu, R.T.
 Direct Submission
 Submitted (30-JUL-1999) Molecular Medicine Center for
 Cardiovascular Disease, Cardiovascular Institute, CAMS & PUMC, 167
 Bei Li Shi Lu, Beijing 100037, P.R. China
 Location/Qualifiers

FEATURES
 source
 1..1278
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /tissue_type="aorta"
 1..1278
 /gene="MST097"
 591..1010
 /gene="MST097"
 /codon_start=1
 /product="MSTP097"
 /protein_id="AA013623.1"
 /db_xref="GI:3338002"
 /translation="MEIPSPDGIICRGQISSGAIAGPSHSVISMNOIHGRYMGHLTRNN
 OGVLGSVAMQAEKFPPEPNRLLLHGFLENVHFAHTSILLSFLVRAGKPYDLQIYP
 QERHSIRVPSGEHVELHLLHLYLQENLSIAALKVI"

ORIGIN

Query Match 36.3%; Score 1132; DB 4; Length 1278;
 Best Local Similarity 96.0%; Pred. No. 1.5e-302; Mismatches 40; Indels 11; Gaps 6;
 Matches 1227; Conservative 0;

1844 CTGGAGAGTGCAAGGCTGACTGACCGTGGCTACTCACATCTCTGCTGCATCAGTCAGC 1903
 Db 1 CTGGAGAGTGCAAGGCTGACTGACCGTGGCTACTCACATCTCTGCTGCATCAGTCAGC 60
 1904 ACTGTGACTCTTTTATAAGTATAGTAAACCAAGAGATCCACACTGTGTCTCTTT 1963
 Db 61 ACTGTGACTCTTTTATAAGTATAGTAAACCAAGAGATCCACACTGTGTCTCTTT 120
 1964 ACAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCCTCCAGAAATTTTGGGCCACCA 2023
 Db 121 ACAAGCTATCAAGTCTCTGAAGATGACCCAACTTGCCTCCAGAAATTTTGGGCCACCA 180
 2024 TTTTGGATTACAGAGTCTCTTCTCTGACTATCTCTCCAGAAATTTTCTTTTGA 2083
 Db 181 TTTTGGATTACAGAGTCTCTTCTCTGACTATCTCTCCAGAAATTTTCTTTTGA 240
 2084 GTACTACTGATTTACATTTGATGGATGCTCTACAGCCTCATGATCTACAGCCTGGAA 2143
 Db 241 GTACTACTGATTTACATTTGATGGATGCTCTACAGCCTCATGATCTACAGCCTGGAA 300
 2144 AGAATATCTACTGTCTGTTTATATATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 2203
 Db 301 AGAATATCTACTGTCTGTTTATATATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 360
 2204 GGTTTAAAGAGTCAAGTATTCGGCTTGATATACCTAGCTCTCTAGGTTATGTTGTTG 2263
 Db 361 GGTTTAAAGAGTCAAGTATTCGGCTTGATATACCTAGCTCTCTAGGTTATGTTGTTG 420
 2264 TAGTGATAGCAACAGGGATCTGTCCAGCGGCTTAATTTTGAAGGCGCTTTAAAT 2323
 Db 421 TAGTGATAGCAACAGGGATCTGTCCAGCGGCTTAATTTTGAAGGCGCTTTAAAT 480
 2324 ATAAATGGGTCAAAATAGAAATGACGATCAGGTGGAAAGGACTCCAAATATCTAGCTTCTC 2383
 Db 481 ATAAATGGGTCAAAATAGAAATGACGATCAGGTGGAAAGGACTCCAAATATCTAGCTTCTC 540
 2384 GATATG-ATTTCATTGACTTAGATCGTGGGCGATCCACCGCTGGTCTCTATGGAGGATAC 2442

Db 541 GATATGAATTCATTGACTTAGATCGTGGGCATNTCCCGCTGTCTCTATGGAGATAC 600
 2443 CTCTCCCTGATGGCAATTAATCAGAGGTGAGATATCTTTCAGGGTTGCTATTGCTGGGCG 2502
 Db 601 CTCTCTCTGATGGCAT--ATCGAGAGGTGAGATATCTTTCAGG--TGCTATTGCTGGG--C 654
 2503 CAGTCACTCTGTGGATCTTCTATGATACAGGATACAGGAACTTATATGGGTCAACCT 2562
 Db 655 CAGTCACTCTGTGATCTCTAT--GAACCCAGATACACGG--ACGTTATATATGGGTCAACCT 710
 2563 GACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAAGTTCCCC 2622
 Db 711 ACCAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAAGTTCCCC 770
 2623 TCTGAACCAAAATCTTACTGCTCTTACATGTTTCTGATGAGATGAGATGTCATTTTGCA 2682
 Db 771 TTTGAACCAAAATCTTACTGCTCTTACATGTTTCTGATGAGATGAGATGTCATTTTGCA 830
 2683 CATACCAATATATTACTGAGTTTCTGATGAGGCTGGAAGCCATATGATTACAGATC 2742
 Db 831 CATACCAATATATTACTGAGTTTCTGATGAGGCTGGAAGCCATATGATTACAGATC 890
 2743 TATCTCAGAGAGACAGACATGAGTTCTCTGAATCGGAGAACATTATGAATCAT 2802
 Db 891 TATCTCAGAGAGACAGACATGAGTTCTCTGAATCGGAGAACATTATGAATCAT 950
 2803 CTTTTCACATACCTTCAAGAAACCTTGGATCAGTATGCTCTTAAAGTATATAA 2862
 Db 951 CTTTTCACATACCTTCAAGAAACCTTGGATCAGTATGCTCTTAAAGTATATAA 1010
 2863 TTTTGACCTGTGTAGAACTCTCTGTATACATGCTGGTATTTAAACAAATGAGGAGTTTA 2922
 Db 1011 TTTTGACCTGTGTAGAACTCTCTGTATACATGCTGGTATTTAAACAAATGAGGAGTTTA 1070
 2923 ATCAACAGAAAAACACAGAAATTCATCATCACTTTTGATACCTGCGATGTAACATCTACTC 2982
 Db 1071 ATCAACAGAAAAACACAGAAATTCATCATCACTTTTGATACCTGCGATGTAACATCTACTC 1130
 2983 CTGAAATAAATGTGGTGCCATGCGAGGGGTCTACCGTTTGTGGTAGTAATCTAATACCTT 3042
 Db 1131 CTGAAATAAATGTGGTGCCATGCGAGGGGTCTACCGTTTGTGGTAGTAATCTAATACCTT 1190
 3043 AACCCACATGCTCAAAATCAAAATGATACATATCTCTGAGAGACCCAGCAATACCAATAAG 3102
 Db 1191 AACCCACATGCTCAAAATCAAAATGATACATATCTCTGAGAGACCCAGCAATACCAATAAG 1250
 3103 AATTACTTAAAAAAA 3120
 Db 1251 AATTACTTAAAAAAA 1268

RESULT 10
 BM557438
 LOCUS
 DEFINITION
 AGENCOURT_6578992 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5466759
 5', mRNA sequence.
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.
 1 (bases 1 to 1042)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 JOURNAL
 COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbe@mail.nih.gov
 Tissue Procurement: DCTD/DTP
 cDNA Library Preparation: Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 plate: LLM1967 row: k column: 16
 High quality sequence stop: 697.
 Location/Qualifiers
 1. .1042
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5466759"
 /tissue_type="amelanotic melanoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 41"
 /note="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGACAGCAG(G). Library constructed by Ling Hong in the
 laboratory of Gerald M. Rubin (University of California,
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
 Superscript II RT (Life Technologies). Note: this is a
 NIH MGC Library."

ORIGIN

Query Match 30.1%; Score 938.2; DB 3; Length 1042;
 Best Local Similarity 98.7%; Pred. No. 1.1e-248;
 Matches 976; Conservative 0; Mismatches 10; Indels 3; Gaps 3;

QY 1723 GAAGTCTTGGCGGCGATGGATCTAATATCCAAAGTTGATGAAGTCAGAGCGCTGGTATAT 1782
 DB 1 GAAGTCTTGGCGGCGGCGATGGATCTAATATCCAAAGTTGATGAAGTCAGAGCGCTGGTATAT 60

QY 1783 TTTTGAAGGACCAAGAGACTCCCTTTAGAGCATCACCTGACGTAGTCAGTACGTAAAT 1842
 DB 61 TTTTGAAGGACCAAGAGACTCCCTTTAGAGCATCACCTGACGTAGTCAGTACGTAAAT 120

QY 1843 CTTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACATTTCTTGTCATCAGTCAG 1902
 DB 121 CTTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACATTTCTTGTCATCAGTCAG 180

QY 1903 CACTGTGACTCTTTTAAGTAAGTATAGTAAACCAAGAGATCCACACTGTGTGCTCCTT 1962
 DB 181 CACTGTGACTCTTTTAAGTAAGTATAGTAAACCAAGAGATCCACACTGTGTGCTCCTT 240

QY 1963 TACAAGTATCAAGTCTCAAGATGCCCACTTCCAAACCAAGAGATTTTGGGCCACC 2022
 DB 241 TACAAGTATCAAGTCTCAAGATGCCCACTTCCAAACCAAGAGATTTTGGGCCACC 300

QY 2023 ATTTTGGATTACGACGCTCTCTCTGACTATACCTCTCCAGAAATTTCTCTTTTGA 2082
 DB 301 ATTTTGGATTACGACGCTCTCTCTGACTATACCTCTCCAGAAATTTCTCTTTTGA 360

QY 2083 AGTACTACTGGATTACATTGTATGGATGCTCTCAAGCCCTCATGATCTACAGCCTGGA 2142
 DB 361 AGTACTACTGGATTACATTGTATGGATGCTCTCAAGCCCTCATGATCTACAGCCTGGA 420

QY 2143 AAGAATAATCTACTGTGCTGTTTATATATGTTGCTCTCAGTGCAGTTGGTGAATAAT 2202
 DB 421 AAGAATAATCTACTGTGCTGTTTATATATGTTGCTCTCAGTGCAGTTGGTGAATAAT 480

QY 2203 CGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATCCCTAGCCTCTCTAGGTTATGCTGT 2262
 DB 481 CGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATCCCTAGCCTCTCTAGGTTATGCTGT 540

QY 2263 GTAGTGATAGACCAAGGGGATCCTGTCAAGGGGCTTAAATTTGAAGGCGCCTTTAAA 2322
 DB 541 GTAGTGATAGACCAAGGGGATCCTGTCAAGGGGCTTAAATTTGAAGGCGCCTTTANA 600

QY 2323 TATAAATGGGTCAAAATAGAAATGACGATCAGGTGGAAGACTCCCAATATCTAGCTTCT 2382
 DB 601 TATAAATGGGTCAAAATAGAAATGACGATCAGGTGGAAGACTCCCAATATCTAGCTTCT 660

QY 2383 CGATATGATTTCATTGACTTAGATCTGTGGGATCCACGGCTGCTCTATGGAGATAC 2442
 DB 661 CGATATGATTTCATTGACTTAGATCTGTGGGATCCACGGCTGCTCTATGGAGATAC 720

QY 2443 CTCTCCCTGATGTCATTAAATGACAGGTCAGATATCTTCCAGGTTGCTATTGCTGGGCC 2502
 DB 721 CTCTCCCTGATGTCATTAAATGACAGGTCAGATATCTTCCAGGTTGCTATTGCTGGGCC 780

QY 2503 CCAGTCACTCTGTGGATCTTCTATATGATACAGGATACCGGAACGTTATATGGTCAACCT 2562
 DB 781 CCAGTCACTCTGTGGATCTTCTATATGATACAGGATACCGGATACGTTATATGGTCAACCT 840

QY 2563 GACAGAATGAACAGGCTATTTACTTAGATCTGTGGCCATCGACGAGAAAGTCCCC 2622
 DB 841 GACAGAATGAACAGGCTATTTACTTTANGATCTGTGGCCATCGACGAGAAAGTCCCC 900

QY 2623 TCTGAACCAATCGTTTACTGCTCTTACATG-GTTTCTGGATGAGATGT-CCATTTTG 2680
 DB 901 TCTGACCAATCGATTACTGCTCTTACATGCTTCTTCTGGATGAGATGTCCCATTTG 960

QY 2681 CACATA-CCAGTATTTACTGAGTTTTTT 2708
 DB 961 CACATACCAAGATATTTACTGGAGTTTTT 989

RESULT 11
 AK050021 3327 bp mRNA linear HTC 03-APR-2004
 LOCUS
 DEFINITION Mus musculus adult male liver tumor cDNA, RIKEN full-length
 enriched library, clone:C730003D12 product:DIPEPTIDYL PEPTIDASE 9
 homolog [Homo sapiens], full insert sequence.
 ACCESSION AK050021
 VERSION AK050021.1 GI:26340743
 KEYWORDS HTC; CAP trapper.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1
 AUTHORS Carninci, P. and Hayashizaki, Y.
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 PUBMED 10349636
 REFERENCE 2
 AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Komoto, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
 PUBMED 11042159
 REFERENCE 3
 AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
 Komoto, H., Akiyama, J., Nishi, K., Kusunagi, T., Tashiro, H., Itoh, M.,
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
 Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 TITLE RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer
 JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
 PUBMED 11076861
 REFERENCE 4
 AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the
 FANTOM Consortium.
 TITLE Functional annotation of a full-length mouse cDNA collection
 JOURNAL Nature 409, 685-690 (2001)
 REFERENCE 5
 AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research
 Group Phase I & II Team.
 TITLE Analysis of the mouse transcriptome based on functional annotation

JOURNAL REFERENCE AUTHORS	of 60,770 full-length cDNAs Nature 420, 563-573 (2002) 6 (bases 1 to 3327) Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Niehi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasakawa, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku, Akahira, S., Takeda, Y., Tanaka, I., Tomari, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
TITLE JOURNAL	Direct Submission Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
COMMENT	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Tissue was provided by William A. Held, Roswell Park Cancer Institute, Department of Molecular and Cellular Biology, Elm and Carlton Streets, Buffalo, NY 14263, whose assistance we gratefully acknowledge. Please visit our web site for further details. URL: http://genome.gsc.riken.jp/ URL: http://fantom.gsc.riken.jp/ Location/Qualifiers
FEATURES	1..3327 ¹ organism="Mus musculus" ¹ mol_type="mRNA" ¹ strain="CE7BL/6J" ¹ db_xref="PANTOM:DB:C730003D12" ¹ db_xref="taxon:10090" ¹ clone="C730003D12" ¹ sex="male" ¹ cell_type="tumor" ¹ tissue_type="liver" ¹ clone_lib="RIKEN full-length enriched mouse cDNA library" ¹ dev_stage="adult" 176..2764 ¹ note="unnamed protein product; DIPEPTIDYL PEPTIDASE 9 homolog [Homo sapiens] (SPTR AAL47179, evidence: FASTA, 92%ID, 99.6%length, match=2580) putative" ¹ codon_start=1 ¹ protein_id="BAC34034.1" ¹ db_xref="GI:26340744" ¹ translation="MCSGVSPVEQVAAGMDMDTAARFCVQKSHWDGLRSIHGSRKSS GLIVKAPHDFOVKPDSPGSHRLYYLGMYPYGSRENSLLYSEIPKVKREALLLL SWKMDHPQATPHGVYSREBELLRKRLGVFTSYDFHSGSLFLFOANSFLH CRDGGKNGFMVPMKPLEIKTCOSGRMDPKICPADPAFESFINNSDLVANIETGEE RLTFCQSSAGVLDNPKSAGVATFVIOEEDRFTGCWCPFTASWEGSEGLKRLILY KEVDESEVIVHPSPALEREKTSYRPTGSKNPKIALKLELQTDHCKIVSSCE KELQVPSFSLPKVEYIAPAGWTRDGVAMFLDRPOQLVLPLPALFIPAVESE ACROAARAVKNNQPVFVIEVNTWINVDIHPFPOAEGQDQFCFLRANECKTF CHLYRVTEKTKVDWTEPLSPTEDEPKIKREVALTSGEMVLSRHGSKIWNVEQ TKLAVYFGKTDPLEHLVYVSEYSGIIVRLITLFGSHCSMSQSPDMFVSHVSVS TPFCVHYLKSDPDDPLHKOPFWASMEANCPDVPVEIPFHPTRDVLGYMI YKPHITLQGRKPTULFYGGVQVOLVNNSPKIKYLPNTLASLGVAVVIVDGRGSC QRGHGFKAIVKMGQVEIDQVGLQVAKYGFIDLSRVAIHGVSYGGLSLMGLI HKQVFKVALAGAPVTVMADITGTERTYMDVPPNNQGVKAGSVLHVSKLPENR LLIHDGFLENVHFHTNLFNLSQLIRAGKPVQLQIYPNERHSIRCSGSEHYEVTLHH FLOEHL"
CDS	
ORIGIN	

QY 1333 GCTTGGTCCATCCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACT 1392
DB 1262 GCCTGGGCGCATGCTTCCTGGGACCGTCCCGCAGCAACGGCTTCCTCTCTCGCCCGCT 1321
QY 1393 GAAATATTTATCCAGTAGAAGATGATGTTATGAAAGGACAGAGACTCATTGAGTCAAGT 1452
DB 1322 GCTCTCTTCATCCCGCGCTTGAGAGTAGGAGCCAGCGGAGCTGCCAGAGCCGTC 1381
QY 1453 CTTGATCTGTGACGCGCACTAATTTATCTATGAAGAAACAAAGACATCTGGATAAATATC 1512
DB 1382 CCCAAGATGTGACGCGCTTGTCTCATCTATGAAGAGTCAACCAATGTCTGGATCAAGTC 1441
QY 1513 CATGACATCTTTCATGTTTTCCTCCAAAGTCA---CGAAGAGAAATTTGATTTATTTT 1569
DB 1442 CACGACATCTTTCACCGCTTTCCTCAGGCTGAGGCGGACGAGCTTTGTTTCTCGT 1501
QY 1570 GCTCTGAATCAAAACAGGTTTCGGTCAATTTATACAAATTTACATCTAATTTAAAGAA 1629
DB 1502 GCCAAGAAATGCAAGACTGGCTTCTGCCACTGTACAGGTCACAGTGAACCTTAAGACC 1561
QY 1630 AGCAATATAAAGCATCCAGTGTGGCTGCTCTCAAGTGATTTCAAGTGTCTTATC 1689
DB 1562 AAGGACTATGACTGACGGAACCCCTCAGCCCTACAGAAGATGATTTAAGTGGCCCATC 1621
QY 1690 AAGAGGAGATAGCAATTTACAGTGTGAATGGGAAGTCTTGGCCGGGATGATCTAAT 1749
DB 1622 AAGGAGGAGTGCCTGACAGTGTGGAGTGGAGGCTCTTTCGAGGCGATGGCTCCAAG 1681
QY 1750 ATCCAGTGTGATGAGTCAAGAGCTGTGATATTTTGAAGGACCAAGAGCTCCCTTTA 1809
DB 1682 ATCTGGGTCAACAGGACAGAGAGCTGTGTACTTTCAAGGTACAAAGSACACACCGCTG 1741
QY 1810 GAGCATCACTGTACGTAGTCACTTAAATCTCTGAGAGGTCACAAAGCTCACTGAC 1869
DB 1742 GAACATCACTCTATGTGTGCTAGCTAGCTAGCAGGCGAGATCGTGGCTCACCAAG 1801
QY 1870 GTGGTACTACATCTTGTGTGATCAGTCACTGTGATCTTTTATAGTAAGTAT 1929
DB 1802 CTGGCTTCTCCACAGCTGTCTCATGAGCCAGAGCTTCGACATGTCTGTGAGTCACTAC 1861
QY 1930 AGTACACAGAGATCCACACTGTGTCTCTTTTACAGCTATCAAGTCTCAAGTCAAGTAC 1989
DB 1862 AGCAGTGTGACACGCGCTGTGTATGTGTACATGTGTACAGTGAAGGCGGCGGATGAC 1921
QY 1990 CCAACTTCCAAAACAAAGGAATTTTGGGCCACCATTTTGGATTTCAAGAGCTCTCTTCT 2049
DB 1922 CCATGACACAGCAACACGCTTCTGGGCCAGCATGATGAGGCGACCAATGCCCCCA 1981
QY 2050 GACTATACCTCCAGAAATTTCTCTTTTGAAGTACTACTGGAATTTACATGTGATGGG 2109
DB 1982 GACTATGTGCCCCCTGAGATCTTCCACTTCCACACCGGTGACAGCTGTACGCGC 2041
QY 2110 ATGCTCTACAGCCTCATGATCTACAGCCTGGAAGAAATATCTACTGTGCTGTCTATA 2169
DB 2042 ATGATCTCAAGGCCACACACCTGCAACCTGGGAGGAACCCCACTGTCTCTTTGTC 2101
QY 2170 TATGTGTGCTCTCAGGTGCAATGATGTAATTCGTTTAAAGGAGTCAAGTATTTCCG 2229
DB 2102 TATGGGGCCCAACAGGTGCGATGTTGTTGAACACTCTTTAAGGCAATCAATACCTGG 2161
QY 2230 TTGAATACCTTAGCCTCTCTAGTGTATGTGTTGTAGTGATAGACACAGGGGATCTGT 2289
DB 2162 CTAATACACTGGCATCTTGGGCTATGCTGTGTGTGATGATGCTGGGGCTCTCTGT 2221
QY 2290 CACGAGGGCTTAAATTTGAAGGGCCCTTTAAATATAAATGGGTCAATAGAAATTCAC 2349
DB 2222 CAGCGGGGCTTGCATCTCGAGGGGGCCCTGAAAAATCAATGGGCGAGTGGAGATTGAG 2281
QY 2350 GATCAGGTGGAGGACTCCATATCTAGCTTCTCGATATGATTTTCAATGACTTAGATCGT 2409
DB 2282 GACCAAGTGGAGGCTTGCAGTACGTGGCTGAGAAGTATGGCTTCAATGACTTGGCCGA 2341
QY 2410 GTGGGCATCCACGGCTGTCTCTATGAGGAGTACCTCTCCCTGATGGCATTAATGACAGG 2469

DB 2342 GTGCCATCCATGGCTGTCTACGGGGCTTCTCTCATCTATGGGGCTATCCACAG 2401
QY 2470 TCAGATATCTTTCAGGGTGTCTATTGTCTGGGGCCCAAGTCACTCTGTGGATCTTCTATGAT 2529
DB 2402 CCACAAGTGTTCAGGTAGCCATTGCGGGCGCTCTGTCTCACTGTGTGGATGGCTTATGAC 2461
QY 2530 ACAGGATACACGGAAGCTTATATGGGTCACTCCCTGACAGAGAAAGTGAACAGGGCTTATCTTA 2589
DB 2462 ACAGGTACACGGAAGCATATCATGGATGTCCTCCGAAATTAACCAAGAGGCTATGAGCA 2521
QY 2590 GATCTGTGGCCATGCAAGCAGAAAAGTTCCTCTGAAACAAATCTGTTTACTGCTCTTA 2649
DB 2522 GAGTCTGTAGCCCTGTGATGTGGAAGAGCTGCCCAATGAGCTTAACCGCTGTATCTCT 2581
QY 2650 CATGTTTCTCTGATGAGAAATGTCATTTTGCACATACCAGTATATTAATCTAGTATTTTA 2709
DB 2582 CAGGCTTCTCGACGAGAGAGTTCCTCTTCTCCACACAAATTTCTGTGTGCCAGCTG 2641
QY 2710 GTGAGGGCTGAAAGCCATATGATTTACAGATCTATCTCTCAGGAGAGACACAGCAAGA 2769
DB 2642 ATCCAGCAGGAAGCCATACCAAGCTTCAGATCTACCAAAACGAGAGACATAGCATCCG 2701
QY 2770 GTTCTGATCGGAGAACATTTATGATGATGATGATGATGATGATGATGATGATGATGAT 2829
DB 2702 TGCCGCGAGTCCGAGAGCATTACGAGGTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2761
QY 2830 GGATCAC 2836
DB 2762 TGACCTC 2768

RESULT 12

AK078301
LOCUS
DEFINITION
Mus musculus adult male olfactory brain cDNA, RIKEN full-length
enriched library, clone:6430584G11 product:DIPEPTIDYL PEPTIDASE 9
homolog [Homo sapiens], full insert sequence.
ACCESSION
VERSION
AK078301.1 GI:26347124
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
10349636
REFERENCE
2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
11042159
REFERENCE
3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kusunagi, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
11076861
REFERENCE
4
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.

Tue Apr 18 08:18:38 2006

Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 3457)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, I.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan [E-mail: genome-res@gsc.riken.jp,
URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216]
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in Riken.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/
Location/Qualifiers
1. 3457
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:6430584G11"
/db_xref="taxon:10090"
/clone="6430584G11"
/sex="male"
/tissue types="olfactory brain"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
182_-2770
/note="unnamed protein product; DIPEPTIDYL PEPTIDASE 9
homolog [Homo sapiens] (SPTR|AAL47179, evidence: FASTA,
9281D_99.6%length, match=2580)
putative"
/codon_start=1
/protein_id="BAC37211.1"
/db_xref="GI:26347125"
/translation="MSGVSPVQVAVAGDMDDTAARFCVQKHSWDGLRSIIHGRKSS
GLVSKAPHDFQVQKPDGSHSLHYLLGMPYGSRENSLLYSEIPKVKRKEALLLL
SWKMLDHFQATPHGVSRBELRLERKRLGVFGITSYDHSRSLGFLFOANSFLPH
CRDGGKFWSPKPLEIKTQCGPRMDPKCPADPAFPSPFINNDSI.WVANIETGSE
RLFTCHQSGAGVLDNPKSGAVATVIOEPRFTGWCWCPATNWSSEGLKTLRLY
BEVDESEVTHVPSPALERKIDSRYRRTGSKPKPKIALKELQTDHQKIVSSCE
KELVQFSSLPFVKEYIARAGNRDGRYANMFLYRFPQAGQDFCFPLANECKTGF
AQRQAARVAKNVPQVYIYEVNVINWVNIHDIYHFRPQAGQDFCFPLANECKTGF
CHLYRVLTCKDYDWTPELSPTEDEFKCPICEVALTSGEVLSRHSKIVNVSQ
TKLVYFOGKDTPLHLLHYVVSYESAGEIVRLITLGFSSCSMSQSPMVFVSHVVS
TPCPVHYKLSGDDPDLHKQPRFMASMEANCPDIYVPEIYFHTFTRAVDQVYGM
YKPHLPQGRKHPTVLVLYGPGQVLNNISFKIGIKYLRNLHLSLGLGYAVVLDGRSC
QRGLHFEGLAKNQMGVEIDQVEGLQYVAEKYFIDLSRVAILHGSVGGELSLMGL
HKPQVFKVIAAGVPTVMAYDTGYTERYMDYVNNKQGTGAGSVALHVEKLPENR
LLILHGFIDENHFFHTNFLVSLRAGKPYQLQIYVNERHRSIRCSEGEHYEVTLLH
FLQEHL"

ORIGIN	Query Match	28.2%	Score 880.6	DB 4	Length 3457
	Best Local Similarity	59.8%	Pred. No. 1.7e-232		
	Matches 1512	Conservative	0	Mismatches 1009	Indels 6
	Gaps				2
QY	316	TTTTATGTTGAGCGGTATTCCTCGAGTCAAGTTAAAGAGCTGCTTGGCGATACACAGAAAA	375		
DB	248	TTCTGTGTGCAGAGCACTCGTGGGATGGGCTGCTAGCAATATCCACGGCAGTGCAG	307		
QY	376	TATCATGGCTACATGATGGTAAAGGCCACCACTAGATTTTCATGTTGTGAGAGGAATCAT	435		
DB	308	TCCTCGGCGCTCATTTGTCAGCAAGGCCCCCAGCACTTCAGTTTGTGCAAGAGCTGAC	367		
QY	436	CCAGATGACCTCAATTCAGACAGAAATCTATTACCTTGGCCATGTCGTGGTGAAGACAGAA	495		
DB	368	GAGTCTGGCCCCCACTCTCCACCGTCTCTATTACCTCGGAATGCTTTACCGCAGCGGTAG	427		
QY	496	AATACACTGTTTATTTCTGAATATCCCAAACTATCAATAGAGACAGACAGTCTTAATGCTC	555		
DB	428	NACTCCCTCTCTACTCCGAGATCCCAAGAAAGTGGGAAGGAGGCCCTGCTGCTGCTG	487		
QY	556	TCCTTGAAGCCCTCTTTTGGATCTTTTTCAGGCAACACTGGACTATCGAATGTATTCTCGA	615		
DB	488	TCCTGGAAGCAGATGCTGGACCACTTCAGGSCCAACACCCACCATGCTGCTACTCCCCA	547		
QY	616	GAAGAGACTATTAAAGAGAAAGAAACGCAATGGAAACAGTCGGAATTTGCTTCTTACGAT	675		
DB	548	GAGGAGGACTACTCGGGAGGCGCAAGCGCTGGGCGTCTTCGGAATCACTCTCTTATGAC	607		
QY	676	TATCACAAGAGTGGGAACATTTCTTCAAGCCGTAGTGGAAATTTATACGTAATAA	735		
DB	608	TTCCAAGTGAAGGCGGCTCTTCTCTTCAGGSCAGCAATACGCTGTTTCCACTGCAGG	667		
QY	736	GATGAGGCGCCACAAGGATTTACGCAACAGCTTTAAGGCCCAATCTAGTGGAAACTAGT	795		
DB	668	GATGTTGGCAAGAAATGGCTTTATGTTGTCCTCCGATGAAGCACTGGAGATCAAGACTCAG	727		
QY	796	TGTCCTCAACATACGAGTGGATCCAAAATATGTCCTCCGCTGATCAGACTGGATTGCTTTT	855		
DB	728	TGTTCTGGGCGCAGCATGGACCCCAAAATCTGCCCCGACACCTGCTGCTTCTTCTTCTTC	787		
QY	856	ATACATAGCAACGATATTTGGATATCTAACATGTAACAGAGAAAGAGAGACTCACT	915		
DB	788	ATCAACAACAGTATCTGTGGGTGGCAACATCGAGACTGGGAGAGAACGGCGCTCACC	847		
QY	916	TATGTGCACAATGAGCTAGCCCAACATGGGAAGAGATGCCAGATCAGCTGGAGTGCCTACC	975		
DB	848	TTCTGTCAACAGGCTTCAAGACGCTGCTGCTGGCAATCCCAATCAGCAGGCGTGGCCACC	907		
QY	976	TTTGTCTTCCAAGAGAAATTTGATAGATATTTCTGGCTATTTGTTGGTGTCCAAAAG---	1032		
DB	908	TTTGTCTTCCAAGAGAGTTCGACCGCTTCACTGGGTGCTGGTGTGCTGCCCGCGCTCT	967		
QY	1032	GAACAACTCCAGTGTGGTAAATTTCTTAGAATTTCTATATGAAGAAATGATGAATCT	1092		
DB	968	TGGGAAGGCTCCGAAGTCTCAAGACGCTGCGCATCTATATAGGAAATGGAAGAGTCT	1027		
QY	1092	GAGTGGAAATTTATGATTTATATCCCTATGTTTGGAAACAGAGAGGCGAGATTCATTC	1152		
DB	1028	GAGTGGAGGTCTATCATGTGCTCCCGCGCTGGAGAGAGGAGAGAGGAGTCTCTAC	1087		
QY	1152	CGTTATCTTAAACAGGATCAGCAATCTTAAGTCACTTTTAAAGATGTCAGAAATAATG	1212		
DB	1088	CGCTACCCAGGACAGGCGCAAGAACCCCAAGATTTGCCCTGAAGCTGGCTGAGCTCCAG	1147		
QY	1212	ATTGATGCTGAAGAGGATCATAGATGTCATAGATAAGAACTAATTCACCTTTTGAG	1272		
DB	1148	ACGGACCATACGGGCAAAATCGTCAAGCTGCGAGAGGAAGTGGTCAAGCATTCAGC	1207		
QY	1272	ATTCTATTGAAGGAGTTGAATATATTGCCAGAGCTGGATGAGTCTCTAGGAGAAATAT	1332		
DB	1208	TCCCTTTTCCCCAAAGTGGAGTACATCGCCCGGCTGGTGGACACGGGACCGCAATAT	1267		

Qy	1333	GCTTGGTCCATCCTACTAGATCGCTCCGAGACTGCGCTACAGATAGTGTGATCTCACCT	1392
Db	1268	GCTTGGGCCATGTTCTCTGTACCGTCCCGCAACGCGCTTACGCTTGCTCTCTACCCCT	1327
Qy	1393	GAATTATTTATCCCACTAGAGATGATGTTTATGGAAGGCACAGACTCATTTAGTTCAGTG	1452
Db	1328	GCTCTCTTCATCCCGCGCGTGTAGAGTGAAGCCACGGGAGCGAGCTGCCMGAGCCGTG	1387
Qy	1453	CCTGATTCTGTGAAGCCACTAAATTTATATGAAGAAACCAACAGACTCTGGATAAATATC	1512
Db	1388	CCCAAGAAATGTGCAGCCCTTTGTTCATCTATGAAGAAAGTCAACCAATGTCTGGATCAA	1447
Qy	1513	CATGACATCTTTTCATGTTTTTTTCCCAAAAGTCA--CGAAGAGAAATTCAGTTTATTTT	1569
Db	1448	CAGACATCTTCCACCCGTTTCTCAGGCTGAGGGCCAGCAGGACTTTTGTTCCTCGT	1507
Qy	1570	GCTCTGTAATGCAAAACAGGTTTCCGTCATTTATACAAAATTACATCTAATTTAAAGGAA	1629
Db	1508	GCCACGAAATGCAGACTGGCTTCTGCACTGTACAGGATCAGAGTGAACACTTAAACC	1567
Qy	1630	AGCAATATPAAACGATCCAGTGTGGGCTCCCTGCTCCAAAGTGATTTCAAGTGTCTATC	1689
Db	1568	AAGGACTATGACTGACGGAACCCCTCAGCCCTACAGAAGATGAGTTTAAAGTGCCTCATC	1627
Qy	1690	AAAGAGGAGATAGCAATTTACAGTGGTGAATGGAAAGTTCTTGGCCGCGATGGATCTAAT	1749
Db	1628	AAAGAGGAGGTGCGCCCTGACCAAGTGGCGATGGGAGGTCTTGTGAGGCGATGGTCCAAG	1687
Qy	1750	ATCCAAGTTGATGAAGTCAGAAAGCTGGTATATTTTGAAGGCACCAAGACTCCCCTTTA	1809
Db	1688	ATCTGGGTCAAGCAGCAGACGAACTGGTGTACTTTCAAGGTACAAAGGACACACCGCTG	1747
Qy	1810	GAGCATACCTGTACGTAGTCAGTTACGTAAATCTCTGGAGAGGTCAAGGCTCACTGAC	1869
Db	1748	GAACATCACCTCTATGTGTCAGCTACAGTCAGCAGCGCAGATCGTGGGCTCACCAAG	1807
Qy	1870	CGTGGCTACTCACATTTCTGTGTCATCAGTCAGCAGTGTGACTTCTTTTATAGTAAAT	1929
Db	1808	CTCGGCTTCTCCACAGCTGTCTCATGAGCCAGAGCTTCGACATGTTGCTGAGTCACTAC	1867
Qy	1930	AGTAAACAGAGAATAACCACACTGTGTCTCCTTTTACAAGCTATCAAGTCTCTGAAGATGAC	1989
Db	1868	AGCAGTGTGACACGCCACCTGTGTACATGTGTACAAGCTGAGCGGCCCGATGATGAC	1927
Qy	1990	CCAACTTGGAAAAACAAAGAAATTTTGGGCCCAACATTTTGGATTTCAGCAGGTCCTCTTC	2049
Db	1928	CCACTGCACAAGCAACCAACGCTTCTGGGGCAGCATGATGGAGGCAGCCAAATGCCCCCA	1987
Qy	2050	GACTATACCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTTACATCTGATGGG	2109
Db	1988	GACTATGTGCCCCCTGAGATCTTCCACTTCCACCCGTCGACAGCGTGAAGCTCTACGGC	2047
Qy	2110	ATGCTCTACAAGCCTCATGATCTACAGCCTGGAAAGAAATATCTCTACTGTGCTGTTCATA	2169
Db	2048	ATGATCTACAGCCACACACCCCTGCAACTCTGGGAGGAGCACCCCACTGTCTCTTTGTC	2107
Qy	2170	TATGTGTGCTCTCAGGTGCAGTTGGTGAATPAATCGGTTTTAAAGAGTCAAGATTTCCGC	2229
Db	2108	TATGGGGGCCCAACAGGTGCAGTTGGTGAACAACCTCTTTAAAGGGCATCAAAATACCTCGG	2167
Qy	2230	TTGAAATACCTTAGCTCTTAGGTTATGTGGTTGTAGTCATAGACAACAGAGGATCTCTGT	2289
Db	2168	CTAAATATACATGGCATCTCTGGGCTATGTGTGGTGGTATCGATGTGCGGGCTCTCTGT	2227
Qy	2290	CACCGAGGCTTTAAATTTGAAGGGCCCTTTTAAATATAAAAATGGGTCAAAATGAAATTTGAC	2349
Db	2228	CAGCGGGCCTGCATCTCGAGGGGGCCCTGAAAAATCAATGGCCAGGTGAGATTGAG	2287
Qy	2350	GATCAGGTGGAGGACTCCAAATATCTAGACTTCTCGATATGATTTTCATTTGACTTATGATCGT	2409
Db	2288	GACCAAGTGGAGGGTTGTCAGTACGTGGCTGAGAAGTATGGCTTCATTTGACTTGAGCCGA	2347

QY	2410	GTGGGATCCACGGCTGGTCTCTATGAGGATACCTCTCCCTGATGGCATTAATGCAGAGG	2469
Db	2348	GTGCGCATCCATGGCTGGTCTCTACGCGGGCTTCTCTCACTCATGGGGCTCATCCCAAG	2407
QY	2470	TCAGATATCTTTCAGGGTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGAT	2529
Db	2408	CCACAGTGTTTCAGGTAGCAATTCGCGGGCGCTCTCTGTCACTGTGTGGATGGCTATGAC	2467
QY	2530	ACAGGATACACGGAAACGTTATATGGGTTCACCCCTGCACAGCAATGAACAGGGCTATTACTTA	2589
Db	2468	ACAGGGTACACGGAACCATACATGATGTCCCCGAAAAATAACAAGCAAGCGCTATGAGCA	2527
QY	2590	GGATCTGTGGCATGCAAGCAGAAAAGTTCCTCTGCAACCAATCGTTTACTGTCTTTA	2649
Db	2528	GGGTCTGTAGCCCTGCATGTGTGAGAAAGCTGCCCAATGAGCCTAACCGCCTGCTTATCCTC	2587
QY	2650	CATGGTTTCCCTGGATGAGATGTCATTTTGGCATACATACAGTATATTACTAGAGTTTTTA	2709
Db	2588	CACGGCTTCTGTGACGAGAACGTTCACTTCTTCCACAAATTTCTGTGTGCCAGCTG	2647
QY	2710	GTGAGGGCTGGAAGCCATATGATTTTACAGATCTATCTCTCAGGAGAGACACAGCATAAAGA	2769
Db	2648	ATCCGAGCAGGAAGCCATACCGACTTCAGATCTTACCCAAACGAGAGACATAGCATCCGC	2707
QY	2770	GTTCTCTGAATCGGGAGAACATTAATGAATCATCATCTTTTGCACCTTCAAGAAAACTTT	2829
Db	2708	TGCCGCGAGTCCGAGAGCATTAACGAGTGAAGTCTGCTGCACCTTCTGACGAGAACCTG	2767
QY	2830	GGATCAC 2836	
Db	2768	TGACCTC 2774	
RESULT 13			
CR985650		884 bp mRNA linear	EST 28-JUN-2005
LOCUS	CR985650 RZPD no.9016 Homo sapiens cDNA clone RZPDp9016N202 5',		
DEFINITION	mRNA sequence.		
ACCESSION	CR985650.1 GI:68279535		
VERSION	CR985650		
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;		
	Hominidae; Homo.		
REFERENCE	1 (bases 1 to 884)		
AUTHORS	Heil, O., Ebert, L., Hennig, S., Henze, S., Radelof, U., Schneider, D. and Korn, B.		
TITLE	Human T-lymphocytes library		
JOURNAL	Unpublished (2005)		
COMMENT	Contact: Inge Arlart		
	RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH		
	Heubnerweg 6, D-14059 Berlin, Germany		
	Email: www.rzpd.de		
	RZPD; RZPDp9016N202.		
	RZPDLIB; (Human T-lymphocytes) RZPD LIB No.9016		
	http://www.rzpd.de/cgi-bin/products/set.cgi?libNo=9016 Contact:		
	Inge Arlart		
	RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH		
	Heubnerweg 6, D-14059 Berlin, Germany		
	Tel: +49 30 32639 100		
	Fax: +49 30 32639 111		
	www.rzpd.de		

FEATURES
SOURCE

/clone="R2PDp9016N202"
/tissue_type="T-Lymphocytes"
/dev_stage="adult"
/lab_host="DH108"
/clone_lib="R2PD no. 9016"
/notes="Vector: POB80LSN_cloned; Site_1: SalI; Site_2:
NotI; Vector:
http://www.rspdp.de/info/vectors/POB80LSN_cloned.pic.shtml
human t-Lymphocytes was prepared from mRNA obtained from
human t-Lymphocytes with a NotI oligo(dT) primer [5',
GACTAGTTCAGATCGGACGCCGCCCTTTTCTTTTCTTTT 3'].
Double-stranded cDNA was ligated to SalI adaptors,
digested with NotI and cloned into the NotI and SalI sites
of the POB80LSN_cloned vector"

ORIGIN

Query Match	28.1%	Score 875.8	DB 7	Length 884
Best Local Similarity	99.3%	Pred. No. 2.2e-231		
Matches 877	Conservative 0	Mismatches 6	Indels 0	Gaps 0
QY 25	GGCGCTGCTACTGCGCGCGTCTTCTTATGTCGCGTTCGCCGCTGGGTGTGACCG 84			
DB 1	GGCGCTGCTACTGCGCGCGTCTTCTTATGTCGCGTTCGCCGCTGGGTGTGACCG 60			
QY 85	GGCGCGCGCGCGAGGACCACTGCACACGACCGAGTGGAGCGGGCGGCGACATGAAG 144			
DB 61	GGCGCGCGCGCGAGGACCACTGCACACGACCGAGTGGAGCGGGCGGCGACATGAAG 120			
QY 145	CGCGCGCGCGCGTCCATAGCGACGTCGGGACGGTCCGGCGGGCGGGGGGAAGGA 204			
DB 121	CGCGCGCGCGCGTCCATAGCGACGTCGGGACGGTCCGGCGGGCGGGGGGAAGGA 180			
QY 205	AAATGCAATGCGCGAGCAATGGAAACAGAACGCTGGGTGTGAGATATTTGAACT 264			
DB 181	AAATGCAATGCGCGAGCAATGGAAACAGAACGCTGGGTGTGAGATATTTGAACT 240			
QY 265	GCGACTGTGAGAGAAATATGAATCAGGATCGGCTAAATGAGCCTTTATGTT 324			
DB 241	GCAGACTGTGAGAGAAATATGAATCAGGATCGGCTAAATGAGCCTTTATGTT 300			
QY 325	GAGCGTATTCCTGGAGTCAGCTTAAAGAGTGTGTCGAGTACAGAAATATCATGCC 384			
DB 301	GAGCGTATTCCTGGAGTCAGCTTAAAGAGTGTGTCGAGTACAGAAATATCATGCC 360			
QY 385	TACATGATGGCTAAGGCACACATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGGA 444			
DB 361	TACATGATGGCTAAGGCACACATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGGA 420			
QY 445	CCTCATTCAGACAGAAATCTATTACCTTGCCATGTCGTGAGAACAGAGAAATACACTG 504			
DB 421	CCTCATTCAGACAGAAATCTATTACCTTGCCATGTCGTGAGAACAGAGAAATACACTG 480			
QY 505	TTTATTCGAAATCCCAAACTATCATAGAGCAGCAGTCTTAATGCTCTCTTGAAG 564			
DB 481	TTTATTCGAAATCCCAAACTATCATAGAGCAGCAGTCTTAATGCTCTCTTGAAG 540			
QY 565	CTCTTTTGGATCTTTTCAGGACACATGGAATGATGATGATTTATCAGTAAAGATGGAGG 624			
DB 541	CTCTTTTGGATCTTTTCAGGACACATGGAATGATGATGATTTATCAGTAAAGATGGAGG 600			
QY 625	CTATTAGAGAAAGAAACGATGGAACAGTCGGAATGCTCTTACGATATACCAA 684			
DB 601	CTATTAGAGAAAGAAACGATGGAACAGTCGGAATGCTCTTACGATATACCAA 660			
QY 685	GGAAAGTGGAAATCTTCTGTTTCAAGCGGTAGTGAATTTATCAGTAAAGATGGAGG 744			
DB 661	GGAAAGTGGAAATCTTCTGTTTCAAGCGGTAGTGAATTTATCAGTAAAGATGGAGG 720			
QY 745	CCACAGGATTTAGCAACCAACCTTTAAGGCCCAATCTAGTGGAAACTAGTGTCCCAAC 804			
DB 721	CCACAGGATTTAGCAACCAACCTTTAAGGCCCAATCTAGTGGAAACTAGTGTCCCAAC 780			
QY 805	ATACGGATGGATCCAAAATTTATGCCCGCTGATCCAGACTGGATTGCTTTTATACATAGC 864			

Db 781	ATACGGATGGATCCAAAATTTATGCCCTGTGATCCANACTGGATTGCTTTTATACATAGC 840	
QY 865	AACGATATTGGATATCTTAACATCTGTAACACAGAGAGAGGA 907	
Db 841	AACGATATTGGATATCTTAACATCTGTAACACANANANAGGA 883	
RESULT 14		
AK050023		
LOCUS	3376 bp mRNA linear HTC 03-APR-2004	
DEFINITION	Mus musculus adult male liver tumor cDNA, RIKEN full-length	
	enriched library, clone:C73003H19 product:DIPEPTIDYL PEPTIDASE 9	
	homolog [Homo sapiens], full insert sequence.	
ACCESSION	AK050023.1 GI:26093756	
VERSION	HTC; CAP trapper.	
KEYWORDS	Mus musculus (house mouse)	
SOURCE	Mus musculus	
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
REFERENCE	1 Carninci, P. and Hayashizaki, Y. High-efficiency full-length cDNA cloning Meth. Enzymol. 303, 19-44 (1999) 10349636	
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.	
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000) 11042159	
JOURNAL	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsu, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamanoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.	
PUBMED	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000) 11076861	
REFERENCE	4 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium. Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001)	
AUTHORS	5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002)	
TITLE	6 (bases 1 to 3376) Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saigo, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.	
JOURNAL	Direct Submision	
REFERENCE	Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome	

Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-resgsc.riken.jp, URL: <http://genome.gsc.riken.jp/>, Tel: 81-45-503-9222).

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Tissue was provided by William A. Held, Roswell Park Cancer Institute, Department of Molecular and Cellular Biology, Elm and Carlton Streets, Buffalo, NY 14263, whose assistance we gratefully acknowledge.

acknowledge:
Please visit our web site for further details.

URL: <http://genome.gsc.riken.jp/>

URL: <http://phantom.gsc.riken.jp/UKT:ncp://genome.gsc.riken.jp/>

FEATURES

source

1. .3376

1. .3376

```

1. .3376
/organism="Mus musculus"

```

```
/organism="MUS MUS"
/mol type="mRNA"
```

```
/mol_type="HRNA"  
/strain="C57BL/6J"
```

```
/db yref="FANTOM DB:C730003H19"
/stain="C5/BL/6J"
```

```
/db_xref="FANTOM_DB:C7"
/db_xref="taxon:70090"
```

/db_xref="taxon:100

```
/clone="C730003H19"
```

```

/sex="male"
/sex="male"

```

```
/cell_type="tumor"
```

```
/tissue_type="liver"
```

```
/clone_lib;
```

```
/dev_stage
```

235. 2824

ORIGIN

[illegible]

Db	1861	GCTGGCTTCTCCACAGCTGCTCCATGAGCCAGAGCTTCGACATGTTGCTGAGTCACTA	1920
Qy	1929	TAGTAACACAGAGAATCCACACTGTGTGCTCTTTTACAAGCTATCAAGTCTCTGAAGATGA	1988
Db	1921	CAGCAGTGTGAGCAGCCACCTGTGTACATGTGTACAAGCTGAGCGGCCGATGATGA	1980
Qy	1989	CCCACTTGCAGAAACAAAGGAATTTTGGGCCACCATTTTGGATTTCAGCAGGTCCTCTTCC	2048
Db	1981	CCCACTGCACAGCAACACAGCTTCTGGCCAGCATGATGGAGGAGCAGCAATTCGCCCCC	2040
Qy	2049	TGACTATACCTCCAGAAATTTCTCTTTTGAAGTACTACTGGAATTTACATCTGATGG	2108
Db	2041	AGCATATGTGCCCTGAGATCTTCCATCTCCACCCGTCGACAGTGCAGCTCTACGG	2100
Qy	2109	GATGCTTACAGCCTCATGATCTACAGCTGGAAGAAATATCTTACTGCTCTTCTAT	2168
Db	2101	CATGATCTACAGGCACACACCTTGCACCTTGGAGGAGCACCACCTGTCTCTTGT	2160
Qy	2169	ATATGGTGTCTCAGGTGACAGTTGGTGAATATCCGTTTAAAGAGTCAAGTATTCGG	2228
Db	2161	CTATGGGGCCACAGGTGCAATTTGGTGAACAACTCTTTAAGGGATCAATATCTGGC	2220
Qy	2229	CTTGAATACCTAGCTCTCTAGGTTATGTGTTGTAGTGTATGATAGACAAACAGGGGATCTCG	2288
Db	2221	GCTAAATACATGGCATCTTGGGCTATGCTGTGGTGTGATCGATGGTGGGGCTCTCG	2280
Qy	2289	TCACCGAGGGCTTAAATTTGAAGGGGCTTTAAATATAAATGGGTCAAATAGAAATGA	2348
Db	2281	TCAGCGGGGCTTGCATCTTCGAGGGGGCTTGAATAATCAAATGGGCAGGTGAGATGA	2340
Qy	2349	CGATCAGGTGGAAGACCTCCAAATATCTAGCTTCTCGATATGATTTCAATGACTTAGATCG	2408
Db	2341	GGACCAAGTGGAGGCTTTCAGTACGTGGCTGAGAAGTATGGCTTCAATGACTTAGAGCG	2400
Qy	2409	TGTGGCATCCACGGCTGCTCTATGGAGGATACCTCTCCCTGATGGCATTAATGACAGAG	2468
Db	2401	AGTCGCATCCATGCTGCTCTACGGGCTCTCTCTCACTCATGSGGCTCATCCACAA	2460
Qy	2469	GTGATATCTTCAGGTTGCTATGCTGGGGCCAGTCACTCTGATCTTCTATCA	2528
Db	2461	GCACCAAGTGTTCAGGTTAGCATTTGGGGGCTCTCTGCTCTGTGTGATGGCTATGA	2520
Qy	2529	TACAGGATACACGGAACTTATATGGGTACCTCTGACAGATGACAGGGCTTACTT	2588
Db	2521	CACAGGGTACACGGAACTATATGATGTCTCCGAAATTAACAGCAGGCTATGAGGC	2580
Qy	2589	AGGATCTGTGGCCATGCAAGCAGAAAAAGTTCCCTCTGAAACCAATCGTTTACTGCTTT	2648
Db	2581	AGGGTCTGTAGCCCTGCAATGTGGAGAAGCTGCGCAATGAGCCTAACCGCCTGTATCTCT	2640
Qy	2649	ACATGGTTTCTGGATGAGAATGTCCATTTTGCACATACCAAGTATATTAAGTGTCTTTT	2708
Db	2641	CCACGGCTTCTGGACGAGAACTTCACTTCTTCCACACAAATTTCTGTTGTCTCCAGCT	2700
Qy	2709	AGTGAGGCTGGAAAGCCATATGATTTACAGATCTATCTCTCAGGAGAGACACAGCATAAG	2768
Db	2701	GATCCGACGAGAAAGCCATACAGCTTTCAGATCTATCCCAAAACAGAGACATAGCATCCG	2760
Qy	2769	AGTTCTGATCGGAGACATTTATGAATCTCATCTTTTGGACATCTCTTCAAGAAACCT	2828
Db	2761	CTGCCGAGTCCGAGACATTAACAGGTGACGCTGCTGACCTTTCTGAGGAACACCT	2820
Qy	2829	TGATATC	2836
Db	2821	GTGACCTC	2828
RESULT 15			
BQ675006			
LOCUS			
DEFINITION			
BQ675006 957 bp mRNA linear EST 15-JUL-2002			
AGENCOURT_8303592 NIH_MGC_102 Homo sapiens cdna clone IMAGE:6274622			
5', mRNA sequence.			

ACCESSION	BQ675006	GI:21785840
VERSION	EST.	
KEYWORDS	Homo sapiens (human)	
SOURCE	Homo sapiens	
ORGANISM	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.	
REFERENCE	1 (bases 1 to 957)	
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/.	
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)	
JOURNAL	Unpublished (1999)	
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgabbs-remail.nih.gov Tissue procurement: ATCC cDNA Library Preparation: Rubin Laboratory DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LCM2456 row: h column: 15 High quality sequence start: 4 High quality sequence stop: 624.	
FEATURES	Location/Qualifiers	
source	1..957	
	/organism="Homo sapiens"	
	/mol_type="mRNA"	
	/db_xref="taxon:9606"	
	/clone="IMAGE:6274622"	
	/tissue_type="epidermoid carcinoma, cell line"	
	/lab_host="DH10B (phage-resistant)"	
	/clone_lib="NIH_MGC_102"	
	/note="Organ: salivary gland; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACAGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."	
ORIGIN		
Query Match	27.2%; Score 848.6; DB 5; Length 957;	
Best Local Similarity	96.9%; Pred. No. 8.3e-224;	
Matches	908; Conservative	0; Mismatches 24; Indels 5; Gaps 4;
Qy	1630	AGCAATATATAACGATCCAGTGTGGGCTGCTGCTCAAGTGATTTCAAGTGTCTATC 1689
Db	9	AGGAATATAACGATCCAGTGTGGGCTGCTGCTCAAGTGATTTCAAGTGTCTATC 68
Qy	1690	AAAGAGGAGTACGATACCAATACCAAGTGTGAATGGGAAGTTCTTGGCCGCGCATGATCTAAT 1749
Db	69	AAAGAGGAGTACCAATACCAAGTGTGAATGGGAAGTTCTTGGCCGCGCATGATCTAAT 128
Qy	1750	ATCAAGTTGATGAAGTCAAGAGGCTGTATATTTGAAGCCACCAAGAGCTCCCTTTA 1809
Db	129	ATCAAGTTGATGAAGTCAAGAGGCTGTATATTTGAAGCCACCAAGAGCTCCCTTTA 188
Qy	1810	GAGCATCACCTGATACCTAGTACCTGATTAATTCCTGGAGAGTGAAGGCTGAC 1869
Db	189	GAGCATCACCTGATACCTAGTACCTGATTAATTCCTGGAGAGTGAAGGCTGAC 248
Qy	1870	CGTGGCTACTCACATTTCTTGTGTCATCAGTACGACTGTGACTTTTATTAAGTAAAT 1929
Db	249	CGTGGCTACTCACATTTCTTGTGTCATCAGTACGACTGTGACTTTTATTAAGTAAAT 308
Qy	1930	AGTAACCAAGAAATCCACACTGTGTCTCTTTTACAAGCTATCAAGTCTCTGAAGATGAC 1989
Db	309	AGTAACCAAGAAATCCACACTGTGTCTCTTTTACAAGCTATCAAGTCTCTGAAGATGAC 368
Qy	1990	CCAACTTGCAGAAACAAAGGAATTTGGGCCACCATTTTGGATTCAGCAGGTCCTCTCT 2049

Db 369 CCAACTTCGCAAAAACAAAGGAATTTTGGGCCACCATTTTGGATTTCAGCAGGTCCTCTTCCT 428
QY 2050 GACTATAGTCTCTCCAGAAAATTTCTCTTTTGAAGTACTACTGGATTTTACATTGTATGGG 2109
Db 429 GACTATAGTCTCTCCAGAAAATTTCTCTTTTGAAGTACTACTGGATTTTACATTGTATGGG 488
QY 2110 ATGCTCTACAAGCCTCATGATCTACAGCCTGGAAAGAAATATCCTACTGTGCTGTTTCA 2169
Db 489 ATGCTCTACAAGCCTCATGATCTACAGCCTGGAAAGAAATATCCTACTGTGCTGTTTCA 548
QY 2170 TATGCTGCTCTCAGGTGCAGTTGGTGAATAATTCGGTTTAAAGGAGTCAAGTATTTCCGC 2229
Db 549 TATGCTGCTCTCAGGTGCAGTTGGTGAATAATTCGGTTTAAAGGAGTCAAGTATTTCCGC 608
QY 2230 TTGAATA-CCCTAGCCTCTCTAGGTATGTGGTTGTAGTAGATAGAC-AACAGGGGATCCT 2287
Db 609 TTGAATACCCCTAGCCTCTCTAGGTATGTGGTTGTAGTAGATAGACAAACAGGGGATCCT 668
QY 2288 GTCAACCGAGGCTTAAATTTGAAGGCGCC--TTTAAATATAAAATGGGTCAAATAGAAAT 2345
Db 669 GTCAACCGAGGCTTAAATTTGAAGGCGCCCTTTTAAATATAAAATGGGTCAAATAGAAAT 728
QY 2346 TGACGATCAGGTGGAAGACTCCAAATATCTAGCTTCTCGATATGATTTTCAATTGACTTAGA 2405
Db 729 TGACGATCAGGGGGAAGACTCCAAATATCTAGCTTCTCGATATGATTTTCAATTGACTTAGA 788
QY 2406 TCGTGTGGGCATCCACGGCTGGTCTATGGAGGATACCTCTCCCTGATGGCATTAATGCA 2465
Db 789 TCGTGTGGGCATCCACGGCTGGTCTATGGAGGATCCCTCTCCCTGAAGGCATTTATGCA 848
QY 2466 GAGGTCAGATATCTTCAGGGTTGCTATTGCTGGGGCCCAAGTCACCTCTG-TGGATCTTCT 2524
Db 849 GAGGTCAGATATCTTCAGGGTTGCTATTGCTGGGGCCCAAGTCACCTCTG-TGGATCTTCT 908
QY 2525 ATGATACAGGATACACGGAACGTTTATATGGGTCACCC 2561
Db 909 ATGATACAGGATACCCGGAACGTTTATATGGGGCCCC 945

Search completed: April 15, 2006, 00:36:30
Job time : 10968 secs

THIS PAGE BLANK (USPTO)

Result No.	Query			ID	Description
	Score	Match	Length		
1	3120	100.0	3120	3	US-10-070-464-2
2	3106.4	99.6	4829	3	US-09-976-674-12
3	2824.4	90.5	4685	3	US-09-976-674-22
4	2806.4	89.9	4676	3	US-09-976-674-20
5	2759.4	88.4	2797	3	US-09-976-594-1103
6	2644.4	84.8	2671	3	US-09-976-674-2
7	2500.4	80.1	4523	3	US-09-976-674-8
8	2056.4	65.9	4309	3	US-09-976-674-14
9	1347	43.2	1669	3	US-10-070-464-6
10	1036	33.2	1197	3	US-10-070-464-4
11	926.4	29.7	1356	3	US-09-976-674-10
12	861	27.6	2617	3	US-09-976-674-4
13	861	27.6	4219	3	US-09-976-674-28
14	861	27.6	4302	3	US-09-976-674-24
15	801.4	25.7	832	3	US-09-976-674-18
16	798.6	25.6	4180	3	US-09-976-674-36
17	798.6	25.6	4263	3	US-09-976-674-34
18	791.2	25.4	1083	3	US-10-070-464-8
19	716.8	23.0	4076	3	US-09-976-674-32
20	716.8	23.0	4159	3	US-09-976-674-30
21	654.4	21.0	4037	3	US-09-976-674-40
22	654.4	21.0	4120	3	US-09-976-674-38
23	633.4	20.3	823	3	US-09-280-115-171
24	472	15.1	620	3	US-09-976-674-16
25	472	15.1	620	3	US-09-976-674-16
26	472	15.1	620	3	US-09-976-674-16
27	472	15.1	620	3	US-09-976-674-16
28	472	15.1	620	3	US-09-976-674-16
29	472	15.1	620	3	US-09-976-674-16
30	472	15.1	620	3	US-09-976-674-16
31	472	15.1	620	3	US-09-976-674-16
32	472	15.1	620	3	US-09-976-674-16
33	472	15.1	620	3	US-09-976-674-16
34	472	15.1	620	3	US-09-976-674-16
35	472	15.1	620	3	US-09-976-674-16
36	472	15.1	620	3	US-09-976-674-16
37	472	15.1	620	3	US-09-976-674-16
38	472	15.1	620	3	US-09-976-674-16
39	472	15.1	620	3	US-09-976-674-16
40	472	15.1	620	3	US-09-976-674-16
41	472	15.1	620	3	US-09-976-674-16
42	472	15.1	620	3	US-09-976-674-16
43	472	15.1	620	3	US-09-976-674-16
44	472	15.1	620	3	US-09-976-674-16
45	472	15.1	620	3	US-09-976-674-16
46	472	15.1	620	3	US-09-976-674-16
47	472	15.1	620	3	US-09-976-674-16
48	472	15.1	620	3	US-09-976-674-16
49	472	15.1	620	3	US-09-976-674-16
50	472	15.1	620	3	US-09-976-674-16
51	472	15.1	620	3	US-09-976-674-16
52	472	15.1	620	3	US-09-976-674-16
53	472	15.1	620	3	US-09-976-674-16
54	472	15.1	620	3	US-09-976-674-16
55	472	15.1	620	3	US-09-976-674-16
56	472	15.1	620	3	US-09-976-674-16
57	472	15.1	620	3	US-09-976-674-16
58	472	15.1	620	3	US-09-976-674-16
59	472	15.1	620	3	US-09-976-674-16
60	472	15.1	620	3	US-09-976-674-16
61	472	15.1	620	3	US-09-976-674-16
62	472	15.1	620	3	US-09-976-674-16
63	472	15.1	620	3	US-09-976-674-16
64	472	15.1	620	3	US-09-976-674-16
65					

Db	241	CTGGGTGTTGAGATATTTGAAACTGGGACTGTGAGAGAAATATTTGAATCACAGGATCGG	300
Qy	301	CCTAAATTTGGAGCCCTTTTATGTTGAGCGGTATTTCTGGAGTCAGCTTTAAAGAGTCGTT	360
Db	301	CCTAAATTTGGAGCCCTTTTATGTTGAGCGGTATTTCTGGAGTCAGCTTTAAAGAGTCGTT	360
Qy	361	GCCGATACCAGAAAATATCATGTGCTACATGATGGCTTAAGGCCACCAATGATTTTCATGTTT	420
Db	361	GCCGATACCAGAAAATATCATGTGCTACATGATGGCTTAAGGCCACCAATGATTTTCATGTTT	420
Qy	421	GTGAAGAGGAATGATCCAGATGGACCTCATTTACAGACAGAAATCTATTAACCTTGCCATGTCT	480
Db	421	GTGAAGAGGAATGATCCAGATGGACCTCATTTACAGACAGAAATCTATTAACCTTGCCATGTCT	480
Qy	481	GGTGAGAACAGAGAAAATACACTGTTTATTTCTGAAATTTCCCAAAATATCAATAGAGCA	540
Db	481	GGTGAGAACAGAGAAAATACACTGTTTATTTCTGAAATTTCCCAAAATATCAATAGAGCA	540
Qy	541	GCAGTCTTAATGCTCTCTTGGAGCCCTTTTGGATCTTTTTCAGGCAACACTGGACTAT	600
Db	541	GCAGTCTTAATGCTCTCTTGGAGCCCTTTTGGATCTTTTTCAGGCAACACTGGACTAT	600
Qy	601	GGATGTATTTCTCGAGAGAGAACTATTAAAGAGAGAAAGCAATTTGGAACAGTCGGA	660
Db	601	GGATGTATTTCTCGAGAGAGAACTATTAAAGAGAGAAAGCAATTTGGAACAGTCGGA	660
Qy	661	ATTGCTCTTACGATTAATACCAAGGAAGTGGACATTTCTGTTTCAAGCCGTTAGTGA	720
Db	661	ATTGCTCTTACGATTAATACCAAGGAAGTGGACATTTCTGTTTCAAGCCGTTAGTGA	720
Qy	721	ATTATATCGTAAAGATGGAGGCCACAAGGATTTTACGCAACACCTTTTAAGGCCAAT	780
Db	721	ATTATATCGTAAAGATGGAGGCCACAAGGATTTTACGCAACACCTTTTAAGGCCAAT	780
Qy	781	CTAGTGGAACTAGTTTGTCCCAACATACGGATGGATCCAAAATTTATGCCCGCTGATCCA	840
Db	781	CTAGTGGAACTAGTTTGTCCCAACATACGGATGGATCCAAAATTTATGCCCGCTGATCCA	840
Qy	841	GACTGGATTTGCTTTTATACATAGCAACGATATTGGATATCTAACTCGTAAACAGAGAA	900
Db	841	GACTGGATTTGCTTTTATACATAGCAACGATATTGGATATCTAACTCGTAAACAGAGAA	900
Qy	901	GAAGGAGACTCACCTTATGTGCACATGAGCTAGCCAACTGGAAGAAATGCCAGATCA	960
Db	901	GAAGGAGACTCACCTTATGTGCACATGAGCTAGCCAACTGGAAGAAATGCCAGATCA	960
Qy	961	GCTGGAGTCGCTACCTTTGTTCTCCAGAGAAATTTGATAGATATTTCTGGCTATTGGTGG	1020
Db	961	GCTGGAGTCGCTACCTTTGTTCTCCAGAGAAATTTGATAGATATTTCTGGCTATTGGTGG	1020
Qy	1021	TGTCAAAAGCTGAAAGCACTCCCAAGTGGTGAATTTCTTAGAATTTCTATAGAGAA	1080
Db	1021	TGTCAAAAGCTGAAAGCACTCCCAAGTGGTGAATTTCTTAGAATTTCTATAGAGAA	1080
Qy	1081	AATGATGAATCTGAGGTGGAAATTTATCATGTTACATCCCTATGTTGGAACAAGGAGG	1140
Db	1081	AATGATGAATCTGAGGTGGAAATTTATCATGTTACATCCCTATGTTGGAACAAGGAGG	1140
Qy	1141	GCAGATTCATCCGTTATCTCAAACAGGTTACAGCAATCTTAAAGTCACTTTTAAGATG	1200
Db	1141	GCAGATTCATCCGTTATCTCAAACAGGTTACAGCAATCTTAAAGTCACTTTTAAGATG	1200
Qy	1201	TCAGAAATTAATGATTGATGCTGAAGAGGATCATAGATCTCATAGATAAGGAATTAAT	1260
Db	1201	TCAGAAATTAATGATTGATGCTGAAGAGGATCATAGATCTCATAGATAAGGAATTAAT	1260
Qy	1261	CAACCTTTTGGATTTCTAATTTGAAGGAGTTGAATATATTGCCAGAGCTGGATGCTCT	1320
Db	1261	CAACCTTTTGGATTTCTAATTTGAAGGAGTTGAATATATTGCCAGAGCTGGATGCTCT	1320
Qy	1321	GAGGAGAAATATGCTTGGTCCATCTACTAGATCGCTCCCAAGCTCGCTACAGATAGTG	1380
Db	1321	GAGGAGAAATATGCTTGGTCCATCTACTAGATCGCTCCCAAGCTCGCTACAGATAGTG	1380

Qy	1381	TTGATCTCACCTGAATTAATTTATCCAGTAGAAGATGATGTTATGAAAGGAGAGACTC	1440
Db	1381	TTGATCTCACCTGAATTAATTTATCCAGTAGAAGATGATGTTATGAAAGGAGAGACTC	1440
Qy	1441	ATTGAGTCAGTGCCTGATTTCTGTGACGCCACTAATTTATCTATGAGAAACACAGACATC	1500
Db	1441	ATTGAGTCAGTGCCTGATTTCTGTGACGCCACTAATTTATCTATGAGAAACACAGACATC	1500
Qy	1501	TGATAAATATCCATGACATCTTTTCATGTTTTCCTCCCAAAAGTCACAGAGGAAATTTGAG	1560
Db	1501	TGATAAATATCCATGACATCTTTTCATGTTTTCCTCCCAAAAGTCACAGAGGAAATTTGAG	1560
Qy	1561	TTTATTTTGGCTCTGAAATGCAAAACAGGTTTCCTCATTTTATACAAAAATACATCTATT	1620
Db	1561	TTTATTTTGGCTCTGAAATGCAAAACAGGTTTCCTCATTTTATACAAAAATACATCTATT	1620
Qy	1621	TTAAAGGAAAGCAAAATATAAATACGATCCAGTGGTGGCTGCTCTCCAAAGTGAATTTCAAG	1680
Db	1621	TTAAAGGAAAGCAAAATATAAATACGATCCAGTGGTGGCTGCTCTCCAAAGTGAATTTCAAG	1680
Qy	1681	TGTCTCTATCAAGAGGAGATAGCAATTAACAGTGGTGAATGGGAAGTTCTTGGCCGGCAT	1740
Db	1681	TGTCTCTATCAAGAGGAGATAGCAATTAACAGTGGTGAATGGGAAGTTCTTGGCCGGCAT	1740
Qy	1741	GGATCTAATATCCAAAGTTGATGAAGTCAGAAAGGCTGGTATATTTTGAAGGCACCAAGAC	1800
Db	1741	GGATCTAATATCCAAAGTTGATGAAGTCAGAAAGGCTGGTATATTTTGAAGGCACCAAGAC	1800
Qy	1801	TCCCTTTTAGAGCATCACCTGTAGTCAGTTTACGTAAATCTCTGGAGAGGTGACAAGG	1860
Db	1801	TCCCTTTTAGAGCATCACCTGTAGTCAGTTTACGTAAATCTCTGGAGAGGTGACAAGG	1860
Qy	1861	CTGACTGACCGTGGCTACTCATCTTCTGTCGATCAGTCAGCACTGTGACTTCTTTATA	1920
Db	1861	CTGACTGACCGTGGCTACTCATCTTCTGTCGATCAGTCAGCACTGTGACTTCTTTATA	1920
Qy	1921	AGTAAATATAGTAAACAGAGAAATCCACATCTGTGTGTCCTTTTACAAAGCTATCAAGTCT	1980
Db	1921	AGTAAATATAGTAAACAGAGAAATCCACATCTGTGTGTCCTTTTACAAAGCTATCAAGTCT	1980
Qy	1981	GAAGATGACCCAACTTGCAAAACAAAGGAAATTTTGGGCCAACCAATTTTGGATTCAGCAGGT	2040
Db	1981	GAAGATGACCCAACTTGCAAAACAAAGGAAATTTTGGGCCAACCAATTTTGGATTCAGCAGGT	2040
Qy	2041	CCTCTTCTGACTATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTTACA	2100
Db	2041	CCTCTTCTGACTATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTTACA	2100
Qy	2101	TTGTATGGGATCTCTACAGCCTCATGATCTACAGCTGGAAGAAATATCTCTACTGTG	2160
Db	2101	TTGTATGGGATCTCTACAGCCTCATGATCTACAGCTGGAAGAAATATCTCTACTGTG	2160
Qy	2161	CTGTTTCATATATGTTGGTCTCCTCAGTGCAGTTGGTGAATTAATCGGTTTAAAGGAGTCAAG	2220
Db	2161	CTGTTTCATATATGTTGGTCTCCTCAGTGCAGTTGGTGAATTAATCGGTTTAAAGGAGTCAAG	2220
Qy	2221	TATTTCCGCTCAATACCTCTAGCTCTCTAGGTTATGTTGGTGTAGTGTAGTACACAGG	2280
Db	2221	TATTTCCGCTCAATACCTCTAGCTCTCTAGGTTATGTTGGTGTAGTGTAGTACACAGG	2280
Qy	2281	GGATCTGTCCAGGCGCTTAAATTTCAAGCGCCCTTAAATATAAATGAGTCAAAATA	2340
Db	2281	GGATCTGTCCAGGCGCTTAAATTTCAAGCGCCCTTAAATATAAATGAGTCAAAATA	2340
Qy	2341	GAAATTTGACGATCAGGTGGAGGACTCCAATATCTAGCTTCTCATATGATTTTCTGAC	2400
Db	2341	GAAATTTGACGATCAGGTGGAGGACTCCAATATCTAGCTTCTCATATGATTTTCTGAC	2400
Qy	2401	TTAGATCGTGTGGGATCCACGGCTGGTCTTATGGAAGATACCTCTCTCCCTGATGGCATTA	2460
Db	2401	TTAGATCGTGTGGGATCCACGGCTGGTCTTATGGAAGATACCTCTCTCCCTGATGGCATTA	2460

QY 2461 ATGCAGAGTCAATATCTTCAGGGTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGATC 2520
DB |||||
QY 2461 ATGCAGAGTCAATATCTTCAGGGTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGATC 2520
DB |||||
QY 2521 TTCTATGATACAGGATACACGGAACGTTATATGGGTCACTGACAGGATGACAGGCG 2580
DB |||||
QY 2521 TTCTATGATACAGGATACACGGAACGTTATATGGGTCACTGACAGGATGACAGGCG 2580
DB |||||
QY 2581 TATTACTTAGGATCTGTGGCCATGCAAGCAGAGAAAGTTCCCTCTGAAACCAATCGTTTA 2640
DB |||||
QY 2581 TATTACTTAGGATCTGTGGCCATGCAAGCAGAGAAAGTTCCCTCTGAAACCAATCGTTTA 2640
DB |||||
QY 2641 CTGCTTACATGTTTCTCGATGAGAAATGTCATTTTGGCAATACCAAGTATTTACTG 2700
DB |||||
QY 2641 CTGCTTACATGTTTCTCGATGAGAAATGTCATTTTGGCAATACCAAGTATTTACTG 2700
DB |||||
QY 2701 AGTTTTTTAGTGGGCTGGAAGCCATATGATTTTACAGATCTATCTCAGAGAGACAC 2760
DB |||||
QY 2701 AGTTTTTTAGTGGGCTGGAAGCCATATGATTTTACAGATCTATCTCAGAGAGACAC 2760
DB |||||
QY 2761 AGCATAAGAGTTCTCGAATCGGAGAACATTTATGAATGCTATCTTTTGCACCTTCAA 2820
DB |||||
QY 2761 AGCATAAGAGTTCTCGAATCGGAGAACATTTATGAATGCTATCTTTTGCACCTTCAA 2820
DB |||||
QY 2821 GAAACCTTTGGATCAGTATTTGCTCTCTAAAGTGATATAATTTTGAAGCTGTGAGAAC 2880
DB |||||
QY 2821 GAAACCTTTGGATCAGTATTTGCTCTCTAAAGTGATATAATTTTGAAGCTGTGAGAAC 2880
DB |||||
QY 2881 TCTCTGGTATACATCTGGCTATTTAAACCAATGAGAGGTTTAAATCAACAGAAACACAGA 2940
DB |||||
QY 2881 TCTCTGGTATACATCTGGCTATTTAAACCAATGAGAGGTTTAAATCAACAGAAACACAGA 2940
DB |||||
QY 2941 ATTGATCATCATTTTGCATCTGCTGCAATGTAACATCTCTCTGAAATAAATGTTGGT 3000
DB |||||
QY 2941 ATTGATCATCATTTTGCATCTGCTGCAATGTAACATCTCTCTGAAATAAATGTTGGT 3000
DB |||||
QY 3001 CCATGAGGGGCTACGGTTTGTGGTAGTAATCTAATACCTTAACCCACATGCTCAAAA 3060
DB |||||
QY 3001 CCATGAGGGGCTACGGTTTGTGGTAGTAATCTAATACCTTAACCCACATGCTCAAAA 3060
DB |||||
QY 3061 TCAATATGATACATATCTCTGAGAGACCCAGCAATACCAATGAGAGGTTTAAATCAACAGAAACACAGA 3120
DB |||||
QY 3061 TCAATATGATACATATCTCTGAGAGACCCAGCAATACCAATGAGAGGTTTAAATCAACAGAAACACAGA 3120
DB |||||

RESULT 2

US-09-976-674-12
; Sequence 12, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akineanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPP1V
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 4829
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-976-674-12

Query Match 99.6%; Score 3106.4; DB 3; Length 4829;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3119; Conservative 0; Mismatches 1; Indels 2; Gaps 1;
QY 1 AAGTGCTAAAGCCTCCGAGGCCAAGGCCGCTGCTACTGCGCGCTGCTTCTTAGTGGCG 60

DB |||||
QY 1 AAGTGCTAAAGCCTCCGAGGCCAAGGCCGCTGCTACTGCGCGCTGCTTCTTAGTGGCG 60
DB |||||
QY 61 CGTTGCGCGCTGGGTGTGTACCGCGCGCGCCGAGGAGCACTGCAACAGGACCG 120
DB |||||
QY 61 CGTTGCGCGCTGGGTGTGTACCGCGCGCGCCGAGGAGCACTGCAACAGGACCG 120
DB |||||
QY 121 GAGTGGAGCGCGGCGAGCATGAGCGGGCGAGGCGCGCTCCNTAGCGCACTGCGGAGCG 180
DB |||||
QY 121 GAGTGGAGCGCGGCGAGCATGAGCGGGCGAGGCGCGCTCCNTAGCGCACTGCGGAGCG 180
DB |||||
QY 181 TCCGCGCGCGCGCGGAGGAGAAATGCAACATGCGAGCAGCAATGGAACAGAAACAGAACAG 240
DB |||||
QY 181 TCCGCGCGCGCGCGGAGGAGAAATGCAACATGCGAGCAGCAATGGAACAGAAACAGAACAG 240
DB |||||
QY 241 CTGGGTGTTGAGATATTTGAAACTGCGGCACTGTGAGGAGAAATATTTGAATCAAGGATCGG 300
DB |||||
QY 241 CTGGGTGTTGAGATATTTGAAACTGCGGCACTGTGAGGAGAAATATTTGAATCAAGGATCGG 300
DB |||||
QY 301 CCTAAATTTGAGCGCTTTTATGTTGAGCGGTATTTCTGGAGTCAGCTTAAAGAGCTGCTT 360
DB |||||
QY 301 CCTAAATTTGAGCGCTTTTATGTTGAGCGGTATTTCTGGAGTCAGCTTAAAGAGCTGCTT 360
DB |||||
QY 361 GCCGATACAGAAATATCATGCGCTACATGATGCTAAGGCAACCAATGATTTTCTATGTTT 420
DB |||||
QY 361 GCCGATACAGAAATATCATGCGCTACATGATGCTAAGGCAACCAATGATTTTCTATGTTT 420
DB |||||
QY 421 GTGAAGAGCAATGATTCAGATGCGACTCATTCAGACAGAAATCTATTACCTTTGCCATGCT 480
DB |||||
QY 421 GTGAAGAGCAATGATTCAGATGCGACTCATTCAGACAGAAATCTATTACCTTTGCCATGCT 480
DB |||||
QY 481 GGTGAGAACAGAGAAATACTGTTTATTTCTGAAATTCCTCAAACTATCAATAGAGCA 540
DB |||||
QY 481 GGTGAGAACAGAGAAATACTGTTTATTTCTGAAATTCCTCAAACTATCAATAGAGCA 540
DB |||||
QY 541 GCAGTCTTAATGCTCTCTTGGAGCGCTCTTTTGGATCTTTTTCAGGCAACACTGAGCTAT 600
DB |||||
QY 541 GCAGTCTTAATGCTCTCTTGGAGCGCTCTTTTGGATCTTTTTCAGGCAACACTGAGCTAT 600
DB |||||
QY 601 GGAATGTTTCTCGAGAGAAAGAACTATTAAAGAGAAAGAAAGAAAGCACTTGGAACTCGGA 660
DB |||||
QY 601 GGAATGTTTCTCGAGAGAAAGAACTATTAAAGAGAAAGAAAGAAAGCACTTGGAACTCGGA 660
DB |||||
QY 661 ATTGCTTCTTACGATTTATCAACAGAGAGTGAACATTTCTGTTTCAAGCGGTAGTGA 720
DB |||||
QY 661 ATTGCTTCTTACGATTTATCAACAGAGAGTGAACATTTCTGTTTCAAGCGGTAGTGA 720
DB |||||
QY 721 ATTTATCACTAAAGATGAGGGCCACAGGATTTTACGCAACCTTTTAAAGSCCAAT 780
DB |||||
QY 721 ATTTATCACTAAAGATGAGGGCCACAGGATTTTACGCAACCTTTTAAAGSCCAAT 780
DB |||||
QY 781 CTAGTGAAACTAGTTGTCTCCCAACATACGATGGATCCAAATATTATGCCCGCTGATCCA 840
DB |||||
QY 781 CTAGTGAAACTAGTTGTCTCCCAACATACGATGGATCCAAATATTATGCCCGCTGATCCA 840
DB |||||
QY 841 GACTGGAATGCTTTTATACATAGCAACGATATTTGGATATCTAATCTGTAACAGAGAA 900
DB |||||
QY 841 GACTGGAATGCTTTTATACATAGCAACGATATTTGGATATCTAATCTGTAACAGAGAA 900
DB |||||
QY 901 GAAAGGAGACTCACTTATGTGCAATGAGCTAGCCCAACNTGGAAGAGATGCCAGATCA 960
DB |||||
QY 901 GAAAGGAGACTCACTTATGTGCAATGAGCTAGCCCAACNTGGAAGAGATGCCAGATCA 960
DB |||||
QY 961 GCTGGAGTCTGCTACCTTTCTCCCAAGAGAAATTTGATAGATTTCTGGCTATTGGTGG 1020
DB |||||
QY 961 GCTGGAGTCTGCTACCTTTCTCCCAAGAGAAATTTGATAGATTTCTGGCTATTGGTGG 1020
DB |||||
QY 1021 TGTCCAAAGCTGAAACAACTCCAGTGGTGTGTAATTTCTTAGAATTTCTATATGAAGAA 1080
DB |||||
QY 1021 TGTCCAAAGCTGAAACAACTCCAGTGGTGTGTAATTTCTTAGAATTTCTATATGAAGAA 1080
DB |||||
QY 1081 AATGATGAATCTGAGTGGAAATTTATTCATGTTACATCCCTTATGTTGGAACAGAGG 1140
DB |||||

Db 1081 AATGATGAACTCTGAGGTGGAAATTAATTCATGTTTACATCCCTATATGTTGGAACAAGGAGG 1140
Qy 1141 GCAGATTCATTCCTGTTATCCTTAAACAGGTAACGAAATCCTTAAAGTCACTTTTAAAGATG 1200
Db 1141 GCAGATTCATTCCTGTTATCCTTAAACAGGTAACGAAATCCTTAAAGTCACTTTTAAAGATG 1200
Qy 1201 TCAGAAATPAATGATGATGCTGAGGAGGATCATAGATGTCATAGATAAGAACTAAAT 1260
Db 1201 TCAGAAATPAATGATGATGCTGAGGAGGATCATAGATGTCATAGATAAGAACTAAAT 1260
Qy 1261 CAACCTTTTGAGATTCCTATTTGAGGAGGTTGAAATATATTTGCCAGAGCTGGATGGAATCCT 1320
Db 1261 CAACCTTTTGAGATTCCTATTTGAGGAGGTTGAAATATATTTGCCAGAGCTGGATGGAATCCT 1320
Qy 1321 GAGGAAATATGCTTGGTCCATCCTACTAGATCGCTCCAGAGCTCGCCCTACAGATAGTG 1380
Db 1321 GAGGAAATATGCTTGGTCCATCCTACTAGATCGCTCCAGAGCTCGCCCTACAGATAGTG 1380
Qy 1381 TTGATCTCACCTGAAATTAATTTATCCCAAGTAGAAGATGATGTTATGGAAGGAGGAGACTC 1440
Db 1381 TTGATCTCACCTGAAATTAATTTATCCCAAGTAGAAGATGATGTTATGGAAGGAGGAGACTC 1440
Qy 1441 ATTGAGTCAGTCCCTGATTCCTGAGCCCACTAAATATCTATGAGAAACACACACATC 1500
Db 1441 ATTGAGTCAGTCCCTGATTCCTGAGCCCACTAAATATCTATGAGAAACACACACATC 1500
Qy 1501 TGGATAATATCCATGACATCTTCAATGTTTCCCAAGTCAAGAGGAGGAAATTCAG 1560
Db 1501 TGGATAATATCCATGACATCTTCAATGTTTCCCAAGTCAAGAGGAGGAAATTCAG 1560
Qy 1561 TTTATTTTGGCTCTGAAATGCAAAACAGGTTTCGGTCAATTTATACAAAATACATCTAAT 1620
Db 1561 TTTATTTTGGCTCTGAAATGCAAAACAGGTTTCGGTCAATTTATACAAAATACATCTAAT 1620
Qy 1621 TTTAAAGGAAACAAATATAACGATCCAGTGGTGGCTGCTGCTCCAAAGTGAATTTCAAG 1680
Db 1621 TTTAAAGGAAACAAATATAACGATCCAGTGGTGGCTGCTGCTCCAAAGTGAATTTCAAG 1680
Qy 1681 TGCTCTATCAAAGAGGAGATAGCAATTTACCAGTGGTGAATGGGAAGTTCTTGGCCGGCAT 1740
Db 1681 TGCTCTATCAAAGAGGAGATAGCAATTTACCAGTGGTGAATGGGAAGTTCTTGGCCGGCAT 1740
Qy 1741 GGATCTAATATCCAAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCCACCAAGAC 1800
Db 1741 GGATCTAATATCCAAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCCACCAAGAC 1800
Qy 1801 TCCCTTTAGAGCATCACTCTGATAGTCAGTTACGTAAATCCTTGAGAGGTGACAAAG 1860
Db 1801 TCCCTTTAGAGCATCACTCTGATAGTCAGTTACGTAAATCCTTGAGAGGTGACAAAG 1860
Qy 1861 CTGACTGACCGTGGCTACTCACATTTCTGCTGATCAGTCAGTCAGTCACTTCTTTATA 1920
Db 1861 CTGACTGACCGTGGCTACTCACATTTCTGCTGATCAGTCAGTCAGTCACTTCTTTATA 1920
Qy 1921 AGTAAGTATAGTAACCAAGAAATCCACACTGTGTGTCCTTTACAGCTATCAAGTCT 1980
Db 1921 AGTAAGTATAGTAACCAAGAAATCCACACTGTGTGTCCTTTACAGCTATCAAGTCT 1980
Qy 1981 GAAGATGACCCAACTTGCAAAACAAAGAAATTTTGGGCCACCAATTTTGGATTCAGAGGT 2040
Db 1981 GAAGATGACCCAACTTGCAAAACAAAGAAATTTTGGGCCACCAATTTTGGATTCAGAGGT 2040
Qy 2041 CCTCTCTGACTACTACTCTCCAGAAATTTTCTCTTTTGAAGTACTCTGATTTACA 2100
Db 2041 CCTCTCTGACTACTACTCTCCAGAAATTTTCTCTTTTGAAGTACTCTGATTTACA 2100
Qy 2101 TTGTATGGGATGCTCTACAAGCCTCATGATCTACAGCCTGGAAAGAAATATCTACTGTG 2160
Db 2101 TTGTATGGGATGCTCTACAAGCCTCATGATCTACAGCCTGGAAAGAAATATCTACTGTG 2160
Qy 2161 CTGTTTCATATATGTTGG--TCCTCAGGTCGATGTTGGTGAATATTCGTTTAAAGGAGTCA 2218
Db 2161 CTGTTTCATATATGTTGGTCTCTCCTCAGGTGCAATGTTGGTGAATATTCGTTTAAAGGAGTCA 2220

Qy 2219 AGTATTTCCGCTTCAATACCTAGCCTCTCTAGTGTATGTTGGTTAGTAGACAACA 2278
Db 2221 AGTATTTCCGCTTCAATACCTAGCCTCTCTAGTGTATGTTGGTTAGTAGACAACA 2280
Qy 2279 GGGGATCTGTCAACGAGGCTTAAATTTGAAGGCCCTTTAAATATAAAATGGGTCAAA 2338
Db 2281 GGGGATCTGTCAACGAGGCTTAAATTTGAAGGCCCTTTAAATATAAAATGGGTCAAA 2340
Qy 2339 TAGAAATTTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATGATTTCAATG 2398
Db 2341 TAGAAATTTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATGATTTCAATG 2400
Qy 2399 ACTTAGATCGTGTGGGCATCCACGGCTGGTCTTATGAGAGGATACCTCTCCCTGATGGCAT 2458
Db 2401 ACTTAGATCGTGTGGGCATCCACGGCTGGTCTTATGAGAGGATACCTCTCCCTGATGGCAT 2460
Qy 2459 TAAATCAGAGGTTCAGATATCTTCAAGGTTTCTATTTGCTGGGGCCCAAGTCTCTGTGGA 2518
Db 2461 TAAATCAGAGGTTCAGATATCTTCAAGGTTTCTATTTGCTGGGGCCCAAGTCTCTGTGGA 2520
Qy 2519 TCTTCTATGATACAGGATACACGGAACGTTTATATGGGTCAACCTTGACCAAGATGAAACAGG 2578
Db 2521 TCTTCTATGATACAGGATACACGGAACGTTTATATGGGTCAACCTTGACCAAGATGAAACAGG 2580
Qy 2579 GCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGAAACCAAAATCGTT 2638
Db 2581 GCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGAAACCAAAATCGTT 2640
Qy 2639 TACTGCTCTTACATGGTTTCTGATGAGAAATGTCCATTTTGACATACACAGTATATTAC 2698
Db 2641 TACTGCTCTTACATGGTTTCTGATGAGAAATGTCCATTTTGACATACACAGTATATTAC 2700
Qy 2699 TGAGTTTTTTAGTGAGGCTGGAAAGCCATATGATTTTACAGATCTATCTCAGGAGAGAC 2758
Db 2701 TGAGTTTTTTAGTGAGGCTGGAAAGCCATATGATTTTACAGATCTATCTCAGGAGAGAC 2760
Qy 2759 ACAGCATAAAGATTCCTGATTCGGGAGAACATTTATGAACTGCACTTTTGCACTACTTTC 2818
Db 2761 ACAGCATAAAGATTCCTGATTCGGGAGAACATTTATGAACTGCACTTTTGCACTACTTTC 2820
Qy 2819 AAGAAAACTTGGATTCAGTATTCCTGCTCTAAAGTGATATAATTTTGAACCTGTGAGA 2878
Db 2821 AAGAAAACTTGGATTCAGTATTCCTGCTCTAAAGTGATATAATTTTGAACCTGTGAGA 2880
Qy 2879 ACTCTCTGGTAPACACTGGCTATTTAAACCAAAATGAGAGGTTTAAATCAACAGAAAAACA 2938
Db 2881 ACTCTCTGGTAPACACTGGCTATTTAAACCAAAATGAGAGGTTTAAATCAACAGAAAAACA 2940
Qy 2939 GAATTGATCATCACATTTTGTATCTGCAATGTAACTCTAATACCTTAAACCCCAATGCTCAA 2998
Db 2941 GAATTGATCATCACATTTTGTATCTGCAATGTAACTCTAATACCTTAAACCCCAATGCTCAA 3000
Qy 2999 TGCCATGCAAGGCTCTACGGTTTGTGGTAGTAAATCTAATACCTTAAACCCCAATGCTCAA 3058
Db 3001 TGCCATGCAAGGCTCTACGGTTTGTGGTAGTAAATCTAATACCTTAAACCCCAATGCTCAA 3060
Qy 3059 AATCAAAATGATACATATTCCTGAGAGACCCAGCAATACCAATAGAAATTTACTAAAAAANA 3118
Db 3061 AATCAAAATGATACATATTCCTGAGAGACCCAGCAATACCAATAGAAATTTACTAAAAAANA 3120
Qy 3119 AA 3120
Db 3121 AA 3122

RESULT 3
US-09-976-674-22
; Sequence 22, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akineanya, Karen

APPLICANT: Riviere, Pierre
APPLICANT: Junien, Jean-Louis
TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPP1V
FILE REFERENCE: 70669
CURRENT APPLICATION NUMBER: US/09/976,674
CURRENT FILING DATE: 2001-10-12
PRIOR APPLICATION NUMBER: US 60/240,117
PRIOR FILING DATE: 2000-10-12
NUMBER OF SEQ ID NOS: 61
SOFTWARE: PatentIn version 3.1
SEQ ID NO 22
LENGTH: 4685
TYPE: DNA
ORGANISM: Homo sapiens
US-09-976-674-22

Query Match 90.5%; Score 2824.4; DB 3; Length 4685;
Best Local Similarity 95.4%; Pred. No. 0;
Matches 2977; Conservative 0; Mismatches 1; Indels 142; Gaps 1;

Qy 1 AAGTGTCTAAAGCCTCCGAGGCCAAGCGCGCTGCTACTGCCCGCGCTGCTTCTTAGTGCCG 60
Db 1 AAGTGTCTAAAGCCTCCGAGGCCAAGCGCGCTGCTACTGCCCGCGCTGCTTCTTAGTGCCG 60

Qy 61 CGTTCGCCGCTGGGTGTTCACCGCGCGCGCGCGAGGAAGCACTGCAACGAGGACG 120
Db 61 CGTTCGCCGCTGGGTGTTCACCGCGCGCGCGCGAGGAAGCACTGCAACGAGGACG 120

Qy 121 GAGTGAGGCGCGGAGCATGAAGCGCGCGAGCGCGCTCCATAGCCAGCTCGGAGCG 180
Db 121 GAGTGAGGCGCGGAGCATGAAGCGCGCGAGCGCGCTCCATAGCCAGCTCGGAGCG 180

Qy 181 TCCGSGCGGCGCGGAGGAAGAAATGCAACATGGCAGCAATGGAAACAGAAACAG 240
Db 181 TCCGSGCGGCGCGGAGGAAGAAATGCAACATGGCAGCAATGGAAACAGAAACAG 240

Qy 241 CTGGGTGTTGAGATATTGAAACTCGGGA CTGTGAGGAGAAATTTGAATCACAGGATCGG 300
Db 241 CTGGGTGTTGAGATATTGAAACTCGGGA CTGTGAGGAGAAATTTGAATCACAGGATCGG 300

Qy 301 CTTAAATTTGAGCGCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAAAGCTGCTT 360
Db 301 CTTAAATTTGAGCGCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAAAGCTGCTT 360

Qy 361 GCCGATACCAAGAAATATCATGCTACATGATGGCTAAGGCACCAATGATTTTCATGTTT 420
Db 361 GCCGATACCAAGAAATATCATGCTACATGATGGCTAAGGCACCAATGATTTTCATGTTT 420

Qy 421 GTCAAGAGGAATGATCCAGATGGA CTTCAATCAGACAGAACTCTATTACCTTGCCATGCT 480
Db 421 GTCAAGAGGAATGATCCAGATGGA CTTCAATCAGACAGAACTCTATTACCTTGCCATGCT 480

Qy 481 GGTGAGAACAGAGAAATACATGTTTATTCGTAATTCCTGAAATTCCTGAAATCTATCAATAGAGCA 540
Db 481 GGTGAGAACAGAGAAATACATGTTTATTCGTAATTCCTGAAATTCCTGAAATCTATCAATAGAGCA 540

Qy 541 GCAGTCTTAATGCTCTCTTGGAGCTCTTTTGGATCTTTTTCAGGCAACACTGGAAT 600
Db 541 GCAGTCTTAATGCTCTCTTGGAGCTCTTTTGGATCTTTTTCAGGCAACACTGGAAT 600

Qy 601 GGAATGTTATCTCGAAGAGAACTATTTAAGAGAAAGAAACGCATTTGAAACAGTCGGA 660
Db 601 GGAATGTTATCTCGAAGAGAACTATTTAAGAGAAAGAAACGCATTTGAAACAGTCGGA 660

Qy 661 ATTGCTCTTTACGATTATCACCAAGGAGTGAAACATTTCTGTTTCAAGCGGTAGTGA 720
Db 661 ATTGCTCTTTACGATTATCACCAAGGAGTGAAACATTTCTGTTTCAAGCGGTAGTGA 720

Qy 721 ATTTATCACGCTAAAGATGAGGGCCACAGGATTTACGCAACACCTTTAAGGCCCAAT 780
Db 721 ATTTATCACGCTAAAGATGAGGGCCACAGGATTTAAGGCCCAACCTTTAAGGCCCAAT 780

Qy 781 CTAGTGAAACTAGTTGTCCTCCACATACGATGGATCCAAATTTATGCCCCGCTGATCCA 840

Db 781 CTAGTGAAACTAGTTGTCCTCCACATACGATGGATCCAAATTTATGCCCTGCTGATCCA 840

Qy 841 GACTGGATTGCTTTTATACATAGCAACGATATTTGGATATCTAACTGCTAAACAGAGAA 900

Db 841 GACTGGATTGCTTTTATACATAGCAACGATATTTGGATATCTAACTGCTAAACAGAGAA 900

Qy 901 GAAAGGAGACTCACTTATGTGCAATGAGCTAGCCAACTGGAAGAAGATGCCAGATCA 960

Db 901 GAAAGGAGACTCACTTATGTGCAATGAGCTAGCCAACTGGAAGAAGATGCCAGATCA 960

Qy 961 GCTGGAGTCGCTACCTTTGCTTCCAAAGAAATTTGATAGATATTTCTGGCTATTGCTGG 1020

Db 961 GCTGGAGTCGCTACCTTTGCTTCCAAAGAAATTTGATAGATATTTCTGGCTATTGCTGG 1020

Qy 1021 TGTCCAAAAGCTGAAACAACTCCAGTGGTGGTAAAAATTTTGAATTTCTATATGAAGAA 1080

Db 1021 TGTCCAAAAGCTGAAACAACTCCAGTGGTGGTAAAAATTTTGAATTTCTATATGAAGAA 1080

Qy 1081 AATGATGAATCTGAGGTGGAATTTATTCATGTTATCATCCCTATGTTGGAAAACAGGAGG 1140

Db 1081 AATGATGAATCTGAGGTGGAATTTATTCATGTTATCATCCCTATGTTGGAAAACAGGAGG 1140

Qy 1141 GCAGATTCATTCCGTTTATCTTAAACAGGTACAGCAATCTTAAAGTCACCTTTTAAAGATG 1200

Db 1141 GCAGATTCATTCCGTTTATCTTAAACAGGTACAGCAATCTTAAAGTCACCTTTTAAAGATG 1200

Qy 1201 TCAGAAATATGATGATGCTGAAGGAAGGATCATAGATGTCTATAGATAAGGAATTAAT 1260

Db 1201 TCAGAAATATGATGATGCTGAAGGAAGGATCATAGATGTCTATAGATAAGGAATTAAT 1260

Qy 1261 CAACCTTTTGAAGATCTTATTTGAAGAGTTGAATATATGTCAGAGCTGGATGAGTCCCT 1320

Db 1261 CAACCTTTTGAAGATCTTATTTGAAGAGTTGAATATATGTCAGAGCTGGATGAGTCCCT 1320

Qy 1321 GAGGGAATATGCTTGGTCCATCTCTAGATGCTCCAGACTCCCTACAGATAGTG 1380

Db 1321 GAGGGAATATGCTTGGTCCATCTCTAGATGCTCCAGACTCCCTACAGATAGTG 1380

Qy 1381 TTGATCTCACCTGAAATTTATTTATCCAGTAGAAGATGATTTATGGAAGAGGACAGACTC 1440

Db 1381 TTGATCTCACCTGAAATTTATTTATCCAGTAGAAGATGATTTATGGAAGAGGACAGACTC 1440

Qy 1441 ATTGAGTCAGTGGCTGATTCTGTGACGCCACTAATTTATCTATGAAGAAACAAAGACATC 1500

Db 1441 ATTGAGTCAGTGGCTGATTCTGTGACGCCACTAATTTATCTATGAAGAAACAAAGACATC 1500

Qy 1501 TGGATAAATTCATGACATCTTTTCATGTTTTCCTCCAAAGTCACGAGAGGAATTTGAG 1560

Db 1501 TGGATAAATTCATGACATCTTTTCATGTTTTCCTCCAAAGTCACGAGAGGAATTTGAG 1560

Qy 1561 TTTATTTTGGCTCTCAATGCAAAACAGGTTTCCGTCATTTATACAAAATTTACATCTATT 1620

Db 1561 TTTATTTTGGCTCTCAATGCAAAACAGGTTTCCGTCATTTATACAAAATTTACATCTATT 1620

Qy 1621 TTTAAAGGAAGCAAAATATAAAGTCAGTGGTGGGCTGCTGCTCCAAAGTGAATTTCAAG 1680

Db 1621 TTTAAAGGAAGCAAAATATAAAGTCAGTGGTGGGCTGCTGCTCCAAAGTGAATTTCAAG 1680

Qy 1681 TGTCTTATCAAGAGAGATAGCAATTCACAGTGGTGAATGGGAAGTTCTTGGCCGCGCAT 1740

Db 1681 TGTCTTATCAAGAGAGATAGCAATTCACAGTGGTGAATGGGAAGTTCTTGGCCGCGCAT 1740

Qy 1741 GGATCTTAATCAAGTGTGATGAAGTCAGAGGCTGTTATTTTGAAGGACCAAGAGAC 1800

Db 1741 GGATCTTAATCAAGTGTGATGAAGTCAGAGGCTGTTATTTTGAAGGACCAAGAGAC 1800

Qy 1801 TCCCTTTAGAGCATCACTGTACGTAGTCACTAGTAAATCTCTGGAAGGTCGACAGG 1860

Db 1801 TCCCTTTAGAGCATCACTGTACGTAGTCACTAGTAAATCTCTGGAAGGTCGACAGG 1860

Qy 1861 CTGATCGACCGTGGCTACTCACAATTTCTGCTGCATCATGTCAGCTGTGATCTTTTATA 1920

Db 1861 CTGACTGACCGTGGCTACTCACATTTCTGTGTCGATCAGTCGAGCAGCTGACTTCTTTATA 1920
QY 1921 AGTAAGTATAGTAACCAAGAAATCCACACTGTGTGCCCTTTACAAAGCTATCAAGTCCCT 1980
Db 1921 AGTAAGTATAGTAACCAAGAAATCCACACTGTGTGCCCTTTACAAAGCTATCAAGTCCCT 1980
QY 1981 GAAAGATGACCCAACTTGCMAAAACAAAGGAAATTTTGGGCCACCAATTTTGGATTTCAGCAGGT 2040
Db 1981 GAAAGATGACCCAACTTGCMAAAACAAAGGAAATTTTGGGCCACCAATTTTGGATTTCAGCAGGT 2034
QY 2041 CCTCTTCTGACTATACCTCTCCAGAAATTTTCTCTTTGAAAGTACTACTCGATTATACA 2100
Db 2035 ----- 2034
QY 2101 TTGTATGGATGCTCTACAAGCCTCATGATCTACAGCCTGGAAAGAAATATCCTACTGTG 2160
Db 2035 ----- 2034
QY 2161 CTGTTTCATATATGTTGGTCTCAGTGCAGTTCGTGAATATCGGTTTAAAGGATCAAG 2220
Db 2035 -----GTCCTCAGGTGCGTTTGTGAATATCGGTTTAAAGGATCAAG 2078
QY 2221 TATTTCCGTTGAATACCTAGCCTCTAGGTTATGTGTTGTAGTATAGACACACAGG 2280
Db 2079 TATTTCCGTTGAATACCTAGCCTCTAGGTTATGTGTTGTAGTATAGACACACAGG 2138
QY 2281 GGATCCTGTCCCGAGGCTTAAATTTGAAGGCGCTTTAAATATATAAATGGGTCAATA 2340
Db 2139 GGATCCTGTCCCGAGGCTTAAATTTGAAGGCGCTTTAAATATATAAATGGGTCAATA 2198
QY 2341 GAAATTTGAGGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTCAATGAC 2400
Db 2199 GAAATTTGAGGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTCAATGAC 2258
QY 2401 TTAGATCGTGTGGGATCCAGGCTGTCTATGGAGGATACCTCTCCCTGTATGGCATTA 2460
Db 2259 TTAGATCGTGTGGGATCCAGGCTGTCTATGGAGGATACCTCTCCCTGTATGGCATTA 2318
QY 2461 ATGCAGGTCAGATATCTTCAGGGTTGCTATTCCTGGGCGCCAGTCACTCTGTGGATC 2520
Db 2319 ATGCAGGTCAGATATCTTCAGGGTTGCTATTCCTGGGCGCCAGTCACTCTGTGGATC 2378
QY 2521 TTCTATGATACAGGATACACGGAGCTTATATGGTCACTCAGCAGATGAACAGGCG 2580
Db 2379 TTCTATGATACAGGATACACGGAGCTTATATGGTCACTCAGCAGATGAACAGGCG 2438
QY 2581 TATTACTTAGGATCTGTGGCCATGCGAGCAGAAAGTTCCCTCTGNAACCAATCGTTTA 2640
Db 2439 TATTACTTAGGATCTGTGGCCATGCGAGCAGAAAGTTCCCTCTGNAACCAATCGTTTA 2498
QY 2641 CTGCTCTTACATGGTTTCTCGATGAGAAATGTCATTTTGCATACCACTATATTACTG 2700
Db 2499 CTGCTCTTACATGGTTTCTCGATGAGAAATGTCATTTTGCATACCACTATATTACTG 2558
QY 2701 AGTTTCTTACATGGTGGAAAGCCATATGATTTTACAGATCTATCTCAGGAGACAC 2760
Db 2559 AGTTTCTTACATGGTGGAAAGCCATATGATTTTACAGATCTATCTCAGGAGACAC 2618
QY 2761 AGCATAGAGTTCTCGAATCGGGAGAACATTAAGAACTGCACTTTTGCACCTACCTTCAA 2820
Db 2619 AGCATAGAGTTCTCGAATCGGGAGAACATTAAGAACTGCACTTTTGCACCTACCTTCAA 2678
QY 2821 GAAAACTTGGATCAGGTATTCCTCTCTAAAGTGAATATTTTGCACCTGTGTAGAAC 2880
Db 2679 GAAAACTTGGATCAGGTATTCCTCTCTAAAGTGAATATTTTGCACCTGTGTAGAAC 2738
QY 2881 TCTCTGTATACACTGGCTATTTTAAACCAATGAGGAGTTTAAATCAACAGAAAAACAGAA 2940
Db 2739 TCTCTGTATACACTGGCTATTTTAAACCAATGAGGAGTTTAAATCAACAGAAAAACAGAA 2798
QY 2941 ATTGATCATCATTGTTGATCTGCAATGTAACATCTACTCTGAAAAATAAATGTGGTG 3000
Db 2799 ATTGATCATCATTGTTGATCTGCAATGTAACATCTACTCTGAAAAATAAATGTGGTG 2858

QY 3001 CCATCAGGCGCTACGGTTGTGTAGTAATCTAATACCTTTAAACCCACATGCTCAAAA 3060
Db 2859 CCATCAGGCGCTACGGTTGTGTAGTAATCTAATACCTTTAAACCCACATGCTCAAAA 2918
QY 3061 TCAATATGATACATATTTCTGAGAGACCCAGCAATACATTAAGAAATTTCTAAAAA 3120
Db 2919 TCAATATGATACATATTTCTGAGAGACCCAGCAATACATTAAGAAATTTCTAAAAA 2978

RESULT 4

US-09-976-674-20 ; Sequence 20, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 20
; LENGTH: 4676
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-976-674-20

Query Match 89.9%; Score 2806.4; DB 3; Length 4676;
Best Local Similarity 95.1%; Pred. No. 0;
Matches 2968; Conservative 0; Mismatches 1; Indels 151; Gaps 1;

QY 1 AAGTGCTAAAGCCTCCGAGGCCAAGCGCGTGTCTACTGCGCGCGCTCTTTAGTCCG 60
Db 1 AAGTGCTAAAGCCTCCGAGGCCAAGCGCGTGTCTACTGCGCGCGCTCTTTAGTCCG 60
QY 61 CGTTCCGCGCTGGGTTGTCAACCGCGCGCGCGCGAGGAGCGCATCTGCAACACAGGACCG 120
Db 61 CGTTCCGCGCTGGGTTGTCAACCGCGCGCGCGCGAGGAGCGCATCTGCAACACAGGACCG 120
QY 121 GAGTGAGGCGCGCGAGCATGAAGCGCGCGCGCGCGCGCGCTCCATAGCGCATCGGAGCGG 180
Db 121 GAGTGAGGCGCGCGAGCATGAAGCGCGCGCGCGCGCGCGCTCCATAGCGCATCGGAGCGG 180
QY 181 TCCGGCG 240
Db 181 TCCGGCG 240
QY 241 CTGGGTTTGAATATTTGAACTCGGCTGTGAGGAGAAATTTGAATCACAGGATCGG 300
Db 241 CTGGGTTTGAATATTTGAACTCGGCTGTGAGGAGAAATTTGAATCACAGGATCGG 300
QY 301 CCTAAATTTGAGCGCTTTTATGTTGAGCGGTATTCCTGGAGTCAAGTTAAAAAGCTGCTT 360
Db 301 CCTAAATTTGAGCGCTTTTATGTTGAGCGGTATTCCTGGAGTCAAGTTAAAAAGCTGCTT 360
QY 361 GCCGATACCAAGAAATATCATGCTACATGCTAGGCGCGCGCGCGCGCGCGCGCGCGCGCG 420
Db 361 GCCGATACCAAGAAATATCATGCTACATGCTAGGCGCGCGCGCGCGCGCGCGCGCGCGCG 420
QY 421 GTGAAGAGGAAATGATCCAGATGGAATCTATTCAGACAGAAATCTATTCATGATTTTCACTCT 480
Db 421 GTGAAGAGGAAATGATCCAGATGGAATCTATTCAGACAGAAATCTATTCATGATTTTCACTCT 480
QY 481 GGTGAGAACAGAGAAAAATACACTGTTTATTTCTGAAATTCGAAAACTATCAATAGAGCA 540
Db 481 GGTGAGAACAGAGAAAAATACACTGTTTATTTCTGAAATTCGAAAACTATCAATAGAGCA 540

Qy 541 GCAATCTTAATGCTCTCTTTGGAGGCTCTTTTGGATCTTTTCAAGCAACACTGGACTAT 600
Db 541 GCAGTCTTAATGCTCTCTTTGGAGGCTCTTTTGGATCTTTTCAAGCAACACTGGACTAT 600
Qy 601 GGAATGTAATCTCGAGAGAGAACTATTTAGAGAGAGAGAAAGCAATGGAGCAAGTCGGA 660
Db 601 GGAATGTAATCTCGAGAGAGAGAACTATTTAGAGAGAGAGAAAGCAATGGAGCAAGTCGGA 660
Qy 661 ATTGCTTCTTACGATTAATCAACAGAGAGAGTGAACATTTCTGTCTTTCAAGCCGGTAGTGA 720
Db 661 ATTGCTTCTTACGATTAATCAACAGAGAGAGTGAACATTTCTGTCTTTCAAGCCGGTAGTGA 720
Qy 721 ATTATCAGCTAAAGAGATGGAGGGCCACAAGATTTAGCGAAACAACTTTAAAGGCCCAAT 780
Db 721 ATTATCAGCTAAAGAGATGGAGGGCCACAAGATTTAGCGAAACAACTTTAAAGGCCCAAT 780
Qy 781 CTAGTGGAAACTAGTTGTCTCCCAACATACGGATGGATCCAAATTTATGCCCTGTGTATCCA 840
Db 781 CTAGTGGAAACTAGTTGTCTCCCAACATACGGATGGATCCAAATTTATGCCCTGTGTATCCA 840
Qy 841 GACTGGATTCCTTTATACATAGCAACAGATATTTGGATATCTAACATCGTAAACAGAGAA 900
Db 841 GACTGGATTCCTTTATACATAGCAACAGATATTTGGATATCTAACATCGTAAACAGAGAA 900
Qy 901 GAAAGGAGACTCACTTATGTGCACAAATGAGCTAGCCAAATGGAAGAGATGCCAGATCA 960
Db 901 GAAAGGAGACTCACTTATGTGCACAAATGAGCTAGCCAAATGGAAGAGATGCCAGATCA 960
Qy 961 GCTGGAGTCGCTACCTTTGTTCTCCAGAGAAATTTGATAGATATTTCTGGCTATTTGGTGG 1020
Db 961 GCTGGAGTCGCTACCTTTGTTCTCCAGAGAAATTTGATAGATATTTCTGGCTATTTGGTGG 1020
Qy 1021 TGTCCTAAAGCTGAAACAACTCCAGTGGTGAATAATCTTAGAATCTTATATGAAGAA 1080
Db 1021 TGTCCTAAAGCTGAAACAACTCCAGTGGTGAATAATCTTAGAATCTTATATGAAGAA 1080
Qy 1081 AATGATGAATCTGAGGTGGAAATTTATCATGTTACATCCCTATGTTGGAAACAAAGGAGG 1140
Db 1081 AATGATGAATCTGAGGTGGAAATTTATCATGTTACATCCCTATGTTGGAAACAAAGGAGG 1140
Qy 1141 GCAGATTCATTCGGTTATCTTAAACAGGTAACAGAAATCTTAAAGTCACTTTTAAAGATG 1200
Db 1141 GCAGATTCATTCGGTTATCTTAAACAGGTAACAGAAATCTTAAAGTCACTTTTAAAGATG 1200
Qy 1201 TCAGAAATATGATGATGCTGAGGAGGATCAGATGTCATAGATAGAGAACTAAT 1260
Db 1201 TCAGAAATATGATGATGCTGAGGAGGATCAGATGTCATAGATAGAGAACTAAT 1260
Qy 1261 CAACCTTTGAGATTCATTTGAGGAGTTGAATATATGCGAGAGCTGGATGCACTCCT 1320
Db 1261 CAACCTTTGAGATTCATTTGAGGAGTTGAATATATGCGAGAGCTGGATGCACTCCT 1320
Qy 1321 GAGGAAATATGCTTGGTCCATCTTACATAGATGCTCCAGACTGCGCTACAGATAGT 1380
Db 1321 GAGGAAATATGCTTGGTCCATCTTACATAGATGCTCCAGACTGCGCTACAGATAGT 1380
Qy 1381 TTGATCTCACTGAAATTTATTTCCAGTAGAGATGATGTTAGAGAGGACAGACTC 1440
Db 1381 TTGATCTCACTGAAATTTATTTCCAGTAGAGATGATGTTAGAGAGGACAGACTC 1440
Qy 1441 ATTGAGTCAGTGGCTGATTTCTGTGAGCGCACTAAATTTATGAGAGAAACAGAGATC 1500
Db 1441 ATTGAGTCAGTGGCTGATTTCTGTGAGCGCACTAAATTTATGAGAGAAACAGAGATC 1500
Qy 1501 TGGATAAATATCCATGACATCTTTCAATGTTTTTCCCAAGTCAAGAGAGAAATGAG 1560
Db 1501 TGGATAAATATCCATGACATCTTTCAATGTTTTTCCCAAGTCAAGAGAGAAATGAG 1560
Qy 1561 TTTATTTTGGCTGCTGAATGCAAAACAGGTTTCCGTCATTTATACAAAATTTACATCTATT 1620
Db 1561 TTTATTTTGGCTGCTGAATGCAAAACAGGTTTCCGTCATTTATACAAAATTTACATCTATT 1620
Qy 1621 TTAAGGNAAGCAAAATATAACGATCCAGTGGTGGGCTGCTGCTCCAAGTGATTTCAAG 1680

Db 1621 TTAAGGNAAGCAAAATATAACGATCCAGTGGTGGGCTGCTGCTCCAAGTGATTTCAAG 1680
Qy 1681 TGTCTCTATCAAGAGAGAGATAGCAATTTACCAGTGGTGAATGGGAAGTTCTTGGCCGCAT 1740
Db 1681 TGTCTCTATCAAGAGAGAGATAGCAATTTACCAGTGGTGAATGGGAAGTTCTTGGCCGCAT 1740
Qy 1741 GGAATCAATATCCAGTTGATGAAGTCAAGAGGCTGGTATATTTTGAAGGACCAAGAGAC 1800
Db 1741 GGAATCAATATCCAGTTGATGAAGTCAAGAGGCTGGTATATTTTGAAGGACCAAGAGAC 1800
Qy 1801 TCCCTTTTAGAGCATCACCTGTACGTAGTACGTAGTAAATCTCGAGAGGTCACAAGG 1860
Db 1801 TCCCTTTTAGAGCATCACCTGTACGTAGTACGTAGTAAATCTCGAGAGGTCACAAGG 1860
Qy 1861 CTGACTGACCGTGGCTACTCACATTTCTGTGTCATCAGTCAAGCTCTTTTATA 1920
Db 1861 CTGACTGACCGTGGCTACTCACATTTCTGTGTCATCAGTCAAGCTCTTTTATA 1920
Qy 1921 AGTAAGTATAGTAACAGAGAAATCCACTGTGTGTCCTTTTACAAGCTATCAAGTCCT 1980
Db 1921 AGTAAGTATAGTAACAGAGAAATCCACTGTGTGTCCTTTTACAAGCTATCAAGTCCT 1980
Qy 1981 GAAGATGACCCAACTTTGCAAAACAAAGGAATTTTGGGCCACATTTTGGATTCAGCAGGT 2040
Db 1981 GAAGATGACCCAACTTTGCAAAACAAAGGAATTTTGGGCCACATTTTGGATTCAGCAGGT 2040
Qy 2041 CCTCTTCTGACTATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGAATTTTACA 2100
Db 2041 CCTCTTCTGACTATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGAATTTTACA 2100
Qy 2101 TTGTATGGGATGCTCTTACAAGCTCATGATCTACAGCTCGAAGAAATATCTTACTGTG 2160
Db 2101 TTGTATGGGATGCTCTTACAAGCTCATGATCTACAGCTCGAAGAAATATCTTACTGTG 2160
Qy 2161 CTGTTTCAATATATGGTGGTCTCCTCAGGTGCAAGTTGGTGAATTAATCGGTTTTAAGGAGTCAAG 2220
Db 2161 CTGTTTCAATATATGGTGGTCTCCTCAGGTGCAAGTTGGTGAATTAATCGGTTTTAAGGAGTCAAG 2220
Qy 2221 TATTTCCGCTTGAATACCCCTAGCCTCTCTAGGTTATGTGTTGTAGTAGTAGACAACAGG 2280
Db 2281 TATTTCCGCTTGAATACCCCTAGCCTCTCTAGGTTATGTGTTGTAGTAGTAGACAACAGG 2280
Qy 2281 GGAATTCGTCAACGAGGCTTAAATTTGAAGGCGCTTTAAATATAAAATGGGTCAAAATA 2340
Db 2281 GGAATTCGTCAACGAGGCTTAAATTTGAAGGCGCTTTAAATATAAAATGGGTCAAAATA 2340
Qy 2341 GAAATTCAGCATCAGGTGGAGAGCTCCAAATATCTAGCTTCTCGATATGATTTCAATTGAC 2400
Db 2341 GAAATTCAGCATCAGGTGGAGAGCTCCAAATATCTAGCTTCTCGATATGATTTCAATTGAC 2400
Qy 2401 TTGATCTGTGGGTCATCCAGGCTGCTCTATGAGGATACCTCTCCCTGATGCAATTA 2460
Db 2401 TTGATCTGTGGGTCATCCAGGCTGCTCTATGAGGATACCTCTCCCTGATGCAATTA 2460
Qy 2461 ATGAGAGAGTCAAGATATCTTACGGTGTGCTATTTGCTGGGGCCCCAGTCACTCTGFGATC 2520
Db 2461 ATGAGAGAGTCAAGATATCTTACGGTGTGCTATTTGCTGGGGCCCCAGTCACTCTGFGATC 2520
Qy 2521 TTCTATGATACAGGATACAGGAACTTATATATGAGGTCACCTGACAGAAATGACAGGCG 2580
Db 2521 TTCTATGATACAGGATACAGGAACTTATATATGAGGTCACCTGACAGAAATGACAGGCG 2580
Qy 2581 TATTTACTTAGGATCTGGCCATGCAAGCAGAAAGTTTCCCTCTGAAACCAATCGTTTA 2640
Db 2581 TATTTACTTAGGATCTGGCCATGCAAGCAGAAAGTTTCCCTCTGAAACCAATCGTTTA 2640
Qy 2641 CTGCTCTTACATGTTTTCTTGGATGAGATGTCATTTTGGACATACCAAGTATATCTG 2700
Db 2641 CTGCTCTTACATGTTTTCTTGGATGAGATGTCATTTTGGACATACCAAGTATATCTG 2700
Qy 2701 AGTTTTTATGAGGCTGGAAAGGCAATGATGATTTACAGATCTATCTCTCAGGAGAGACAC 2760
Db 2701 AGTTTTTATGAGGCTGGAAAGGCAATGATGATTTACAGATCTATCTCTCAGGAGAGACAC 2760

Db	2550	AGTTTTTTAGTGAGGGCTGGAAAGCCCATATGATTTACAGATCTATCCTCAGAGAGACAC	2609	Db	241	ATCTTTTTCAGGCAACACTGGAATGTATTTCTCGAGAAGAAGAACTATTAAAGAG	300
Qy	2761	AGCATAGAGTTCCTGGAATCGGAGAACATATTAGAACTGCATCTTTTGCACTACCTTCAA	2820	Qy	635	AAAGAAACGATTTGGAAACAGTGGAAATTTCTTACGATTATCACCAGGAAGTGGAA	694
Db	2610	AGCATAGAGTTCCTGGAATCGGAGAACATATTAGAACTGCATCTTTTGCACTACCTTCAA	2669	Db	301	AAAGAAACGATTTGGAAACAGTGGAAATTTCTTACGATTATCACCAGGAAGTGGAA	360
Qy	2821	GAAGAACCTTGGATCACGATTTGCTGCTCTAAAAGTGATATAAATTTTGACCTGTGTAGAAC	2880	Qy	695	CATTTCTGTTTCAAGCCGGTAGTGGAAATTTATCAGCTAAAGATGGAGGCCACAGGAT	754
Db	2670	GAAGAACCTTGGATCACGATTTGCTGCTCTAAAAGTGATATAAATTTTGACCTGTGTAGAAC	2729	Db	361	CATTTCTGTTTCAAGCCGGTAGTGGAAATTTATCAGCTAAAGATGGAGGCCACAGGAT	420
Qy	2881	TCTCTGGGTATACACTGGCTATTTAAACCAATGAGGAGGTTTAAATCAACAGAAAAACACAGA	2940	Qy	755	TTACGCAACAACTTTTAAGGCCCAATCTAGTGAATAGTGTGTCCTCCACATACCGATGG	814
Db	2730	TCTCTGGGTATACACTGGCTATTTAAACCAATGAGGAGGTTTAAATCAACAGAAAAACACAGA	2789	Db	421	TTACGCAACAACTTTTAAGGCCCAATCTAGTGAATAGTGTGTCCTCCACATACCGATGG	480
Qy	2941	ATTGATCATCATTTTGATACCTGCCATGTAACATCTACTCTGAAAAATAAATGTGGTG	3000	Qy	815	ATCCAAAAATATGCCCCGCTGATCCAGACTGGATTCGTTTTATACATAGCAACGATATTT	874
Db	2790	ATTGATCATCATTTTGATACCTGCCATGTAACATCTACTCTGAAAAATAAATGTGGTG	2849	Db	481	ATCCAAAAATATGCCCCGCTGATCCAGACTGGATTCGTTTTATACATAGCAACGATATTT	540
Qy	3001	CCATCAGGGGTCTACGGTTTGTGTAGTAATCTTAATACCTTAACCCACATGCTCAAAA	3060	Qy	875	GGATATCTTAACATCGTAAACCCAGAGAGAAAGGAGACTCACTTATGTGCAAAATGAGCTAG	934
Db	2850	CCATCAGGGGTCTACGGTTTGTGTAGTAATCTTAATACCTTAACCCACATGCTCAAAA	2909	Db	541	GGATATCTTAACATCGTAAACCCAGAGAGAAAGGAGACTCACTTATGTGCAAAATGAGCTAG	600
Qy	3061	TCAATATGATACATATTCCTGAGAGACCCAGCATACCATAGAAATTAACAAAAA	3120	Qy	935	CCAAATCGAAAGAGATGCCAGATCAGCTGGAGTTCGCTACCTTTTGTCTTCCAAAGAGAAAT	994
Db	2910	TCAATATGATACATATTCCTGAGAGACCCAGCATACCATAGAAATTAACAAAAA	2969	Db	601	CCAAATCGAAAGAGATGCCAGATCAGCTGGAGTTCGCTACCTTTTGTCTTCCAAAGAGAAAT	660
RESULT 5							
US-09-976-594-1103							
; Sequence 1103, Application US/09976594							
; Patent No. 6673549							
; GENERAL INFORMATION:							
; APPLICANT: Buchbinder, Jenny							
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS							
; FILE REFERENCE: PA-0041 US							
; CURRENT APPLICATION NUMBER: US/09/976.594							
; PRIORITY FILING DATE: 2001-10-12							
; PRIOR FILING DATE: 2000-10-12							
; NUMBER OF SEQ ID NOS: 1143							
; SOFTWARE: PERL Program							
; SEQ ID NO 1103							
; LENGTH: 2797							
; TYPE: DNA							
; ORGANISM: Homo sapiens							
; FEATURE: misc feature							
; OTHER INFORMATION: Incyte ID No. 6673549 977951.1							
US-09-976-594-1103							
Query Match 88.4%; Score 2759.4; DB 3; Length 2797;							
Best Local Similarity 99.8%; Pred. No. 0;							
Matches 2785; Conservative 0; Mismatches 1; Indels 5; Gaps 2;							
Qy	335	CCTGGAGTCAGCTTAAAAGCTGCTTGGCGATACAGAAAATATCATGGCTACATGATGG	394	Qy	1294	TATATTGCCAGAGCTGGATGGAATCTCTGAGGGAAAAATA-----TGCTTGGTCCATCTCTACT	1349
Db	1	CCTGGAGTCAGCTTAAAAGCTGCTTGGCGATACAGAAAATATCATGGCTACATGATGG	60	Db	961	TATATTGCCAGAGCTGGATGGAATCTCTGAGGGAAAAATAGTGTCTTGGTCCATCTCTACT	1020
Qy	395	CTAAGGCCACCATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTCAG	454	Qy	1350	AGATCGCTCCAGACTCGCCTACAGATAGTGTTCATCTCACCTGAATTTATTTATCCAGT	1409
Db	61	CTAAGGCCACCATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTCAG	120	Db	1021	AGATCGCTCCAGACTCGCCTACAGATAGTGTTCATCTCACCTGAATTTATTTATCCAGT	1080
Qy	455	ACAGAAATCTATTACCTTGGCATGCTCTGGTGAGAACAGAGAAAATACACTGTTTATTCG	514	Qy	1410	AGAAATGATGTTTATGGAAAAGCAGAGACTCAATTGAGTTCAGTGCCTGATTTCTGTGACGCC	1469
Db	121	ACAGAAATCTATTACCTTGGCATGCTCTGGTGAGAACAGAGAAAATACACTGTTTATTCG	180	Db	1081	AGAAATGATGTTTATGGAAAAGCAGAGACTCAATTGAGTTCAGTGCCTGATTTCTGTGACGCC	1140
Qy	515	AAATTCCTCAAACTATCAATAGACGAGCTTTAAATGCTCTCTTGGAAAGCCCTCTTTTGG	574	Qy	1470	ACTAATTAATCTATGAAGAAACAAACAGACATCTGGATAAATATCCATGACATCTTTTATGT	1529
Db	181	AAATTCCTCAAACTATCAATAGACGAGCTTTAAATGCTCTCTTGGAAAGCCCTCTTTTGG	240	Db	1141	ACTAATTAATCTATGAAGAAACAAACAGACATCTGGATAAATATCCATGACATCTTTTATGT	1200
Qy	575	ATCTTTTTCAGGCAACACTGGAATGTATTTCTCGAGAGAGAAAGAACTATTAAAG	634	Qy	1530	TTTTCCCAAGCTCAGAGAGGAATTTAGTGTATTTTTCCTCTCAATGCAAAACAGG	1589
				Db	1201	TTTTCCCAAGCTCAGAGAGGAATTTAGTGTATTTTTCCTCTCAATGCAAAACAGG	1260

```
QY 1710 CAGTGGTGAATGGGAAGTTCTTGGCGGCATGGATCTAATATCCAAAGTTGATGAAGTCAG 1769
Db 1381 CAGTGGTGAATGGGAAGTTCTTGGCGGCATGGATCTAATATCCAAAGTTGATGAAGTCAG 1440
QY 1770 AAGGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACTGTAGCTAGT 1829
Db 1441 AAGGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACTGTAGCTAGT 1500
QY 1830 CAGTTACGTAAATCTCGAGAGGTGACAAAGGCTGACTGACCGGTGCTACTCACAATCTTTG 1889
Db 1501 CAGTTACGTAAATCTCGAGAGGTGACAAAGGCTGACTGACCGGTGCTACTCACAATCTTTG 1560
QY 1890 CTGCATCAGTCAGCACTGTGACTCTTTTAAAGTAAGTATAGTAAACAGAGAAATCCACA 1949
Db 1561 CTGCATCAGTCAGCACTGTGACTCTTTTAAAGTAAGTATAGTAAACAGAGAAATCCACA 1620
QY 1950 CTGTGTGCTCTTTACAGCTATCAAGTCTCTGAAGATGACCAACTTGCARACAAAGGA 2009
Db 1621 CTGTGTGCTCTTTACAGCTATCAAGTCTCTGAAGATGACCAACTTGCARACAAAGGA 1680
QY 2010 ATTTTGGGCGACATTTTGGGATTCAGCAGGTCCTCTTCTGACTATATCTCTCCAGAAAT 2069
Db 1681 ATTTTGGGCGACATTTTGGGATTCAGCAGGTCCTCTTCTGACTATATCTCTCCAGAAAT 1740
QY 2070 TTTCTCTTTTGAAGTACTACTCGGATTTACATTTGATGGATGCTCTACAAGCCTCATGA 2129
Db 1741 TTTCTCTTTTGAAGTACTACTCGGATTTACATTTGATGGATGCTCTACAAGCCTCATGA 1800
QY 2130 TCTACAGCTGAAAGAAATATCTTACTGTGCTGCTTATATATGTTGGTCTCTCAGGTGCA 2189
Db 1801 TCTACAGCTGAAAGAAATATCTTACTGTGCTGCTTATATATGTTGGTCTCTCAGGTGCA 1860
QY 2190 GTTGTGTAATATCGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCTGACCTGCTCT 2249
Db 1861 GTTGTGTAATATCGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCTGACCTGCTCT 1920
QY 2250 AGGTTATGTGGTTGTAGTGATAGACAAACAGGGGATCTGTCAACGAGGGCTTAAATTTGA 2309
Db 1921 AGGTTATGTGGTTGTAGTGATAGACAAACAGGGGATCTGTCAACGAGGGCTTAAATTTGA 1980
QY 2310 AGGCGCTTTAAATATAAAATGGGTCAAATAGAAATTTGACGATCAGGTGAAGGACTCCA 2369
Db 1981 AGGCGCTTTAAATATAAAATGGGTCAAATAGAAATTTGACGATCAGGTGAAGGACTCCA 2040
QY 2370 ATATCTAGCTTCTCGATATGATTTTCAATGATCGTGGGATCCAGGCTGGTC 2429
Db 2041 ATATCTAGCTTCTCGATATGATTTTCAATGATCGTGGGATCCAGGCTGGTC 2100
QY 2430 CTATGAGGATACCTCTCCCTGATGGCAATTAATCAGAGGTCAGATATCTTCAGGGTTGC 2489
Db 2101 CTATGAGGATACCTCTCCCTGATGGCAATTAATCAGAGGTCAGATATCTTCAGGGTTGC 2160
QY 2490 TATTCGTGGGCCCCCAGTCACTGTGTGATCTTCTATGATACAGGATACACGGAACGTTA 2549
Db 2161 TATTCGTGGGCCCCCAGTCACTGTGTGATCTTCTATGATACAGGATACACGGAACGTTA 2220
QY 2550 TATGGGTCACTTGCACAGATGACAGGCTATTTAGTATGATCTGTGGCCTGCAAGC 2609
Db 2221 TATGGGTCACTTGCACAGATGACAGGCTATTTAGTATGATCTGTGGCCTGCAAGC 2280
QY 2610 AGAAAGTTTCCCTCTGACCAAAATCGTTTACTGCTCTTACATGGTTTCTCGATGAGAA 2669
Db 2281 AGAAAGTTTCCCTCTGACCAAAATCGTTTACTGCTCTTACATGGTTTCTCGATGAGAA 2340
QY 2670 TGTCCAATTTTGACATACCAAGTATATTAATGATGATGATGATGATGATGATGATGATGAT 2729
Db 2341 TGTCCAATTTTGACATACCAAGTATATTAATGATGATGATGATGATGATGATGATGATGAT 2400
QY 2730 TGAATTTACAGATCTATCTCAGGAGACACAGCATTAAGATTTCTGAAATCGGGGAAACA 2789
Db 2401 TGAATTTACAGATCTATCTCAGGAGACACAGCATTAAGATTTCTGAAATCGGGGAAACA 2460
```

```
QY 2790 TTATGAATGTCATCTTTTGGCACTACTTCAAGAAAACCTTGGATACGTAATGCTGCTCT 2849
Db 2461 TTATGAATGTCATCTTTTGGCACTACTTCAAGAAAACCTTGGATACGTAATGCTGCTCT 2520
QY 2850 AAAAGTGATATATTTTGAACCTGTGTAGAACTCTCTGTGTATACACTGGCTATTTAAACAA 2909
Db 2521 AAAAGTGATATATTTTGAACCTGTGTAGAACTCTCTGTGTATACACTGGCTATTTAAACAA 2580
QY 2910 ATGAGGAGGTTTAAATCAACAGAAAACACAGAAATGATCATCATATTTTGATACCTGCCAT 2969
Db 2581 ATGAGGAGGTTTAAATCAACAGAAAACACAGAAATGATCATCATATTTTGATACCTGCCAT 2640
QY 2970 GTAACATCTACTCTCTGAAAATAAATGTGTGCCATGCAAGGGTCTACGGTTTGTGTAGT 3029
Db 2641 GTAACATCTACTCTCTGAAAATAAATGTGTGCCATGCAAGGGTCTACGGTTTGTGTAGT 2700
QY 3030 AATCTAATACCTTAACCCACATGCTCAAAATCAAAATGATACATATTTCTTGAGAGACCCA 3089
Db 2701 AATCTAATACCTTAACCCACATGCTCAAAATCAAAATGATACATATTTCTTGAGAGACCCA 2760
QY 3090 GCAATACCATTAAGAATTTACTAAAAAATAA 3120
Db 2761 GCAATACCATTAAGAATTTACTAAAAAATAA 2791

RESULT 6
US-09-976-674-2
; Sequence 2, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 2671
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-976-674-2

Query Match 84.8%; Score 2644.4; DB 3; Length 2671;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2648; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 209 GCACATGCGCAGCAGCAATGGAAACAGAACAGCTGGGTGTGAGATATTTGAAATCGCG 268
Db 3 GTACCATGGCAGCAGCAATGGAAACAGAACAGCTGGGTGTGAGATATTTGAAATCGCG 62
QY 269 ACTGTGAGGAGATATTTGAATCAGAGGATCGGCTTAAATTTGGAGCTTTTATGTTGAGC 328
Db 63 ACTGTGAGGAGATATTTGAATCAGAGGATCGGCTTAAATTTGGAGCTTTTATGTTGAGC 122
QY 329 GGTATTCCTGGAGTCAGCTTAAAGAGCTGCTTCCGATACCAAGAAATATCATGCTGCTACA 388
Db 123 GGTATTCCTGGAGTCAGCTTAAAGAGCTGCTTCCGATACCAAGAAATATCATGCTGCTACA 182
QY 389 TGATGGCTAAGGCACACATGATTTTCATGTTTGTGAAGCAATGATCCAGATGACCTC 448
Db 183 TGATGGCTAAGGCACACATGATTTTCATGTTTGTGAAGCAATGATCCAGATGACCTC 242
QY 449 ATTTCAGACAGAAATCTATTACCTTGGCATGCTGTGTGAGAACAGAGAAAATACACTGTTTT 508
Db 243 ATTTCAGACAGAAATCTATTACCTTGGCATGCTGTGTGAGAACAGAGAAAATACACTGTTTT 302
QY 509 ATTCTGAAATTCCTCAAAACTATCAATAGACGACGAGTCTTAAATGCTCTCTTTGGAAGCCTC 569
```



```

1561 TTTATTTTGGCTCTGAATGCAAAACAGGTTTCGTCATATTTATACAAAATACATCATTT 1620
1621 TTAAGGAAGCAAAATATAAACGATCCAGTGGTGGGCTGCTGCTCCAAAGTGATTTCAAG 1680
1621 TTAAGGAAGCAAAATATAAACGATCCAGTGGTGGGCTGCTGCTCCAAAGTGATTTCAAG 1680
1681 TGTCTCTATCAAAAGAGGAGATAGCAAAATACCAAGTGGTGAATGGGAAGTTCTTGGCCGGCAT 1740
1681 TGTCTCTATCAAAAGAGGAGATAGCAAAATACCAAGTGGTGAATGGGAAGTTCTTGGCCGGCAT 1740
1741 GGATCTAATATCCAAGTTGATGAAGTCAGAAAGCTGGTATATTTTGAAGGCACCAAGAC 1800
1741 GGATCTAATATCCAAGTTGATGAAGTCAGAAAGCTGGTATATTTTGAAGGCACCAAGAC 1800
1801 TCCCTTTTAGAGCATCACTGTACGTAGTCAGTTACGTAAATCTCTGGAGAGGTGACAAGG 1860
1801 TCCCTTTTAGAGCATCACTGTACGTAGTCAGTTACGTAAATCTCTGGAGAGGTGACAAGG 1860
1861 CTGACTGACCGTGGCTACTCACATTTCTTGTGTCATCAGTCAGCAGCTGTGACTTCTTTATA 1920
1861 CTGACTGACCGTGGCTACTCACATTTCTTGTGTCATCAGTCAGCAGCTGTGACTTCTTTATA 1920
1921 AGTAAGTATAGTAACCAAGAGATCCACACTGTGTCTTGTGTCATCAGTCAGCAGCTGTGACT 1980
1921 AGTAAGTATAGTAACCAAGAGATCCACACTGTGTCTTGTGTCATCAGTCAGCAGCTGTGACT 1980
1981 GAAGATGACCAACTGTCGAAACAAAGGAAATTTTGGGCCACCAATTTTGGATTCAGCAGGT 2040
1981 GAAGATGACCAACTGTCGAAACAAAGGAAATTTTGGGCCACCAATTTTGGATTCAGCAGGT 2040
2041 CCTCTTCTGACTACTCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTCGATTTTACA 2100
2041 CCTCTTCTGACTACTCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTCGATTTTACA 2100
2101 TTGTATGGGATGCTCTCAAGCCTCATGATCTACAGCCTGGAAGAAATATCTCTACTGTG 2160
2101 TTGTATGGGATGCTCTCAAGCCTCATGATCTACAGCCTGGAAGAAATATCTCTACTGTG 2160
2161 CTGTTTCATATATGGTGTCCTCAGTGCAGTTGGTGAATAATCGGTTTAAAGGAGTCAAG 2220
2161 CTGTTTCATATATGGTGTC----- 2179
2221 TATTTCCGCTTGAAATACCTAGCCTCTCTAGGTTATGTGTTTGTAGTGATAGACAACAGG 2280
2180 ----- 2179
2281 GGATCCTGTCACCGAGGCTTAAATTTGAAGGCGCCTTTAAATATAAATGGGTCAAATA 2340
2180 ----- 2179
2341 GAAATTGAGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTCTATTGAC 2400
2180 ----- 2179
2401 TTAGATCGTGTGGGCATCCACGGCTGGTGCCTTAGGAGGATACCTCTCCCTGATGGCATTA 2460
2180 ----- 2179
2461 ATGCAGAGTCCAGATATCTTCAGGGTTGCTATTGCTGGGGCCCGAGTCACTCTGTGGATC 2520
2180 -----GGTTGCTATTGCTGGGGCCCGAGTCACTCTGTGGATC 2216
2521 TTCTATGATACAGGATACACGGAACGTTATATGGGTCACTGACCAAGATGAACAGGGC 2580
2217 TTCTATGATACAGGATACACGGAACGTTATATGGGTCACTGACCAAGATGAACAGGGC 2276
2581 TATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGAACCAATCGTTTA 2640
2277 TATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGAACCAATCGTTTA 2336
2641 CTGCTCTTACATGTTTCTCTGATGAGATGTCATTTTGCATACCAAGTATTTACTG 2700
|||||

```

```

2337 CTGCTCTTACATGGTTTCTCGGATGAGAAATGTCANTTTTCACATACCAAGTATATCTG 2396
2701 AGTTTTTTAGTGGGCTGGAAAGCCATATGATTTACAGATCTATCTCTCAGGAGACAC 2760
2397 AGTTTTTTAGTGGGCTGGAAAGCCATATGATTTTACAGATCTATCTCAGGAGACAC 2456
2761 AGCATAGAAGTTCCTGAAATCGGAGAACATTTATGAATGCAATGTCATCTTTTGCNCTACCTTCAA 2820
2457 AGCATAGAAGTTCCTGAAATCGGAGAACATTTATGAATGCAATGTCATCTTTTGCNCTACCTTCAA 2516
2821 GAAAACTTGGATCACGCTATTGCTCTCTAAAAAGTATATAATTTTGAACCTGFTGAGAAC 2880
2517 GAAAACTTGGATCACGCTATTGCTCTCTAAAAAGTATATAATTTTGAACCTGFTGAGAAC 2576
2881 TCTCTGGTATACACTGGCTATTATAACCAAAATGAGGAGGTTTAAATCAACAGAAAACACAGA 2940
2577 TCTCTGGTATACACTGGCTATTATAACCAAAATGAGGAGGTTTAAATCAACAGAAAACACAGA 2636
2941 ATTGATCATACATTTTGTATACCTGTCATCTAACTGTAACATCTACTCTGAAAAATATAATGTTGGTG 3000
2637 ATTGATCATACATTTTGTATACCTGTCATCTAACTGTAACATCTACTCTGAAAAATATAATGTTGGTG 2696
3001 CCATCAGGGGTCTACGGTTTGTGTAGTAAATCTAAATACCTTAAACCCACATGCTCAAAA 3060
2697 CCATCAGGGGTCTACGGTTTGTGTAGTAAATCTAAATACCTTAAACCCACATGCTCAAAA 2756
3061 TCAAATGATACATATTCCTGAGAGACCCAGCAATACCATAGAAATTAATAAAAAA 3120
2757 TCAAATGATACATATTCCTGAGAGACCCAGCAATACCATAGAAATTAATAAAAAA 2816

RESULT 8
US-09-976-674-14
; Sequence 14, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akineanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPP1V
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-976-674-14

Query Match      65.9%; Score 2056.4; DB 3; Length 4309;
Best Local Similarity 83.2%; Pred No. 0;
Matches 2599; Conservative 0; Mismatches 1; Indels 522; Gaps 2;

Qy 1 AAGTCTAAAGCTCTCGAGGCCAGGCGCTGCTACTGCGCGCGCTCTTCTTAGTGCCG 60
Db 1 AAGTCTAAAGCTCTCGAGGCCAGGCGCTGCTACTGCGCGCGCTCTTCTTAGTGCCG 60
Qy 61 CGTTCCGCCCTGGGTTGTFCACCGCGCGCGCGCGAGAGGACCACTGCAACCCAGGACCG 120
Db 61 CGTTCCGCCCTGGGTTGTFCACCGCGCGCGCGCGAGAGGACCACTGCAACCCAGGACCG 120
Qy 121 GAGTGGAGCGCGCAGCATGAAGCGCGCGCGCGCTCCATAGCGCATGTCGGGACGG 180
Db 121 GAGTGGAGCGCGCAGCATGAAGCGCGCGCGCGCTCCATAGCGCATGTCGGGACGG 180
Qy 181 TCCGGCGCGCGCGCGGAGAAAATGCAACATGCGAGCAGCAATGGAAACAGAACAG 240
Db 181 TCCGGCGCGCGCGCGGAGAAAATGCAACATGCGAGCAGCAATGGAAACAGAACAG 240

```

Qy	241	CTGGGTGTTGAGATATTGAACTGCGGACTGTGAGGAGAAATATTGAATCACAGGATCGG	300
Db	241	CTGGGTGTTGAGATATTGAACTGCGGACTGTGAGGAGAAATATTGAATCACAGGATCGG	300
Qy	301	CCTAAATTCGAGCCCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGAGCTGCTT	360
Db	301	CCTAAATTCGAGCCCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGAGCTGCTT	360
Qy	361	GCCGATACACAGAAAATATCATGGCTACATGATGGCTAAGGCACCAATGATTTTCATGTTT	420
Db	361	GCCGATACACAGAAAATATCATGGCTACATGATGGCTAAGGCACCAATGATTTTCATGTTT	420
Qy	421	GTGAAGAGAAATGATCCAGATGAGACCTCATTCAGACAGAAATCTATTACCTTGCATGCT	480
Db	421	GTGAAGAGAAATGATCCAGATGAGACCTCATTCAGACAGAAATCTATTACCTTGCATGCT	480
Qy	481	GGTGAGAACAGAGAAAATACACTGTTTATTCGAAATTCCTCAAACTCAATCAATAGACGA	540
Db	481	GGTGAGAACAGAGAAAATACACTGTTTATTCGAAATTCCTCAAACTCAATCAATAGACGA	540
Qy	541	GCACTCTTAATGCTCTCTTGGAGGCTCTTTTGGATCTTTTTCAGGCAACACTTGGACTAT	600
Db	541	GCACTCTTAATGCTCTCTTGGAGGCTCTTTTGGATCTTTTTCAGGCAACACTTGGACTAT	600
Qy	601	GGAAATGATTTCTCGAAGAGAACTATTAAGAGAGAAAGAAACGCAATGGAAACAGTCGGA	660
Db	601	GGAAATGATTTCTCGAAGAGAACTATTAAGAGAGAAAGAAACGCAATGGAAACAGTCGGA	660
Qy	661	ATTGCTTCTTACGATTTATCACCAAGGAAAGTGGAAACATTTCTGTTTCAAGCCGTTAGTGA	720
Db	661	ATTGCTTCTTACGATTTATCACCAAGGAAAGTGGAAACATTTCTGTTTCAAGCCGTTAGTGA	720
Qy	721	ATTATTCACGTAAGATGGAGGCGCAACAGAAATTTAGCGMAAACCTTTAAGGCCCAAT	780
Db	721	ATTATTCACGTAAGATGGAGGCGCAACAGAAATTTAGCGMAAACCTTTAAGGCCCAAT	780
Qy	781	CTAGTGGAAACTAGTTGTCCTCCAAACATACGAGTGGATCCAAAATTTATGCCCTGTGATCCA	840
Db	781	CTAGTGGAAACTAGTTGTCCTCCAAACATACGAGTGGATCCAAAATTTATGCCCTGTGATCCA	840
Qy	841	GACTGATGCTTTTATACATAGCAACGATATTTGGATATCTAACATCGTAAACAGAGAA	900
Db	841	GACTGATGCTTTTATACATAGCAACGATATTTGGATATCTAACATCGTAAACAGAGAA	900
Qy	901	GAAAGAGACTCACTTATGTGCAACATGAGCTAGCAACATGGAAGAGATGCCAGATCA	960
Db	901	GAAAGAGAGACTCACTTATGTGCAACATGAGCTAGCAACATGGAAGAGATGCCAGATCA	960
Qy	961	GCTGGAGTCGCTACCTTTGTTCTCCAAGAGAAATTTGATAGATATTTCTGGCTATTGGTGG	1020
Db	961	GCTGGAGTCGCTACCTTTGTTCTCCAAGAGAAATTTGATAGATATTTCTGGCTATTGGTGG	1020
Qy	1021	TGTCAAAAGCTGAAACAACTCCAGTGGTGGTAAATTTTGAATTTCTATATGAGAA	1080
Db	1021	TGTCAAAAGCTGAAACAACTCCAGTGGTGGTAAATTTTGAATTTCTATATGAGAA	1080
Qy	1081	AATGATGAATCTGAGGTGAAATTTATTCATGTTATCATCCCTATTTGGTAAACAGAGAGG	1140
Db	1081	AATGATGAATCTGAGGTGAAATTTATTCATGTTATCATCCCTATTTGGTAAACAGAGAGG	1140
Qy	1141	GCAGATTCATTCGGTTATCCTAAACAGGTACAGCAAAATCCTAAGTCACTTTTAAGATG	1200
Db	1141	GCAGATTCATTCGGTTATCCTAAACAGGTACAGCAAAATCCTAAGTCACTTTTAAGATG	1200
Qy	1201	TCAGAAATATGATTCATGCTGAAGAGATCATAGATGTCTATAGATATAGGAACATAAT	1260
Db	1201	TCAGAAATATGATTCATGCTGAAGAGATCATAGATGTCTATAGATATAGGAACATAAT	1260
Qy	1261	CAACCTTTTGAGATTTCTATTGGAAGAGTGTGAATATATTGCCAGAGCTGGATGACTCT	1320
Db	1261	CAACCTTTTGAGATTTCTATTGGAAGAGTGTGAATATATTGCCAGAGCTGGATGACTCT	1320
Qy	1230	-----	1230
Qy	1321	GAGGGAATAATATGCTTGGTCCATCTACTAGATCGTCCCAGACTCGCCTACAGATAGTG	1380
Db	1321	-----	1380
Qy	1381	TTGATCTCACCTGAATTTATTTATCCAGTAGAAGATGATGTTATGGAAGGCAGAGACTC	1440
Db	1381	-----	1440
Qy	1441	ATTGAGTCAGTCGCTGATTCGTGACGCCACTAATTTATCTATATGAAGAAACAAAGACATC	1500
Db	1441	-----	1500
Qy	1501	TGATATAATATCCATGACATCTTTTCATGTTTTTCCCAAAGTCCAGAGAGGAAATTTGAG	1560
Db	1501	-----	1560
Qy	1561	TTTTATTTTGGCTTGAATGCAAAACAGGTTTTCGGTCAATTTATACAAAATTTACATCTATT	1620
Db	1561	-----	1620
Qy	1621	TTAAAGGAAGCAAAATATAAACGATCCAGTGGTGGGCTGCCTGCTCCAGTGATTTCAAG	1680
Db	1621	-----	1680
Qy	1681	TGTCCTTCAAGAGGAGATACCAATTACCAGTGGTGAATGGGAAGTTCTTGGCCGCGAT	1740
Db	1681	-----	1740
Qy	1741	GGATCTAATATCCAGTTGATGAGTCCAGAGGCTGGTATATTTTGAAGGCACCAAGAGAC	1800
Db	1741	-----	1800
Qy	1801	TCCTCTTTAGAGCATCACCTGTACGTAGTCAGTTACGTAAATCTCTGGAGAGGTGCAAGG	1860
Db	1801	-----	1860
Qy	1861	CTGACTGACCGTGGCTACTCACATTTCTGCTGCATCAGTCAGCACTGTGACTCTTTTATA	1920
Db	1861	-----	1920
Qy	1921	AGTAAGTATAGTAAACAGAGAAATCCACATCTGTGTCTCTTTTACAGCTATCAAGTCCT	1980
Db	1921	-----	1980
Qy	1981	GAAAGTGAACCAACTTTGCAAAAACAAAGGAATTTTGGGCCAACATTTTGGATTCAGCAGGT	2040
Db	1981	-----	2040
Qy	2041	CCTCTTCTGACTATCTACTCTCCAGAAATTTTCTCTTTTGAAGTACTACTTGGATTTTACA	2100
Db	2041	-----	2100
Qy	2101	TTGTATGGAGTCTCTACAGGCTCATGATCTACAGCTCGAAGCAATATCTCTACTGTG	2160
Db	2101	-----	2160
Qy	2161	CTGTTTCATATATGGTGG--TCCTCAGGTGCACTTGGTGAATAATCGGTTTAAAGAGGTCA	2218
Db	2161	-----	2218
Qy	2218	AGTATTTTCGGCTTGAATACCTTAGCCTCTCTAGGTTATGTGGTGTAGTATAGACAACA	2278
Db	2218	-----	2278
Qy	2278	GGGATCCTGTCAACGAGGCTTTAAATTTGAAGGGCCCTTTAAATATATAAATGGGTCAAA	2338
Db	2278	-----	2338
Qy	2338	TAGAAATTTGAGATCAGGTGGAGGACTCCAAATATCTAGCTTCTCGATATGATTTCAATTG	2398
Db	2338	-----	2398
Qy	2398	ACTTAGATCGTGTGGGCAATCCAGGCTGGTCTCTATGGAGGATACCTCTCCCTGATGGCAT	2458
Db	2398	-----	2458

```
Db 1881 ACTTAGATCGTGTGGGATCCAGGGCTGGTCTTATGGAGGATACCTCTCCCTGATGGCAT 1940
Qy 2459 TAATGTCAGAGGTGAGATATCTTCAGGGTGTCTATTGCTGGGGCCCGCAGTCACTCTGTGGA 2518
Db 1941 TAATGTCAGAGGTGAGATATCTTCAGGGTGTCTATTGCTGGGGCCCGCAGTCACTCTGTGGA 2000
Qy 2519 TCTTCTATGATACAGGATACACGGACGTTATATGGGTCACTGACGAGAAATGAACAGG 2578
Db 2001 TCTTCTATGATACAGGATACACGGACGTTATATGGGTCACTGACGAGAAATGAACAGG 2060
Qy 2579 GCTATTACTTAGGATCTGTGGCATCAAGCAGAAAGTTCCCTCTGAAACCAATCGTT 2638
Db 2061 GCTATTACTTAGGATCTGTGGCATCAAGCAGAAAGTTCCCTCTGAAACCAATCGTT 2120
Qy 2639 TACTGCTCTTACATGTTTCTGTGAGAGAAATGTCATTTTGCACATACCACTATATTAC 2698
Db 2121 TACTGCTCTTACATGTTTCTGTGAGAGAAATGTCATTTTGCACATACCACTATATTAC 2180
Qy 2699 TGAGTTTTTTAGTGAGGGCTGGAAAGCCATATATGATTTACAGATCTATCTCAGAGAGAC 2758
Db 2181 TGAGTTTTTTAGTGAGGGCTGGAAAGCCATATATGATTTACAGATCTATCTCAGAGAGAC 2240
Qy 2759 ACAGCATAGAGTTCCTGAATCGGGAGAACATTTAGCACTGCACTTTTGGCACTACCTTC 2818
Db 2241 ACAGCATAGAGTTCCTGAATCGGGAGAACATTTAGCACTGCACTTTTGGCACTACCTTC 2300
Qy 2819 AAGAAAACCTTGGATCAGTATCTGCTCTAAAGTGATATAATTTTGACCTGTGTAGA 2878
Db 2301 AAGAAAACCTTGGATCAGTATCTGCTCTAAAGTGATATAATTTTGACCTGTGTAGA 2360
Qy 2879 ACTCTCTGGTATACACTGGCTATTTAAACCAATGAGGAGGTTTAAATCAACAGAAAACCA 2938
Db 2361 ACTCTCTGGTATACACTGGCTATTTAAACCAATGAGGAGGTTTAAATCAACAGAAAACCA 2420
Qy 2939 GAATTTGATCATCACATTTTGATACCTGCGATGAAACATCTACTCTTGAAATAAATGTGG 2998
Db 2421 GAATTTGATCATCACATTTTGATACCTGCGATGAAACATCTACTCTTGAAATAAATGTGG 2480
Qy 2999 TGCCATGAGGGGTCTACGGTTTGTGTAGTATCTAATCTTAATACCTTTAACCCACATGCTCAA 3058
Db 2481 TGCCATGAGGGGTCTACGGTTTGTGTAGTATCTAATCTTAATACCTTTAACCCACATGCTCAA 2540
Qy 3059 AATCAATGATACATATCTCTGAGAGACCCAGCAATACCAATACCAATTAAGAAATTAACAAAAA 3118
Db 2541 AATCAATGATACATATCTCTGAGAGACCCAGCAATACCAATACCAATTAAGAAATTAACAAAAA 2600
Qy 3119 AA 3120
Db 2601 AA 2602
```

```
RESULT 9
US-10-070-464-6
; Sequence 6, Application US/10070464
; Patent No. 6881564
; GENERAL INFORMATION:
; APPLICANT: ABBOTT, Catherine Anne
; APPLICANT: GORRELL, Mark Douglas
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES
; FILE REFERENCE: GH-007
; CURRENT APPLICATION NUMBER: US/10/070.464
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: PCT/AU00/01085
; PRIOR FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: AU PQ5709
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: AU PQ2762
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1669
```

```
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-070-464-6
Query Match 43.2%; Score 1347; DB 3; Length 1669;
Best Local Similarity 84.7%; Pred. No. 0;
Matches 1657; Conservative 0; Mismatches 0; Indels 300; Gaps 1;
Qy 1164 AACAGGTACAGCAAAATCCTAAAGTCACTTTTAAAGATGTCAGAAATTAATGATGATGCTGA 1223
Db 1 AACAGGTACAGCAAAATCCTAAAGTCACTTTTAAAGATGTCAGAAATTAATGATGATGCTGA 60
Qy 1224 AGGAAGGATCATAGATGTCATAGATAAGGAATTAATTAACCTTTTGAAGATCTATTGA 1283
Db 61 AGGAAGGATCATAGATGTCATAGATAAGGAATTAATTAACCTTTTGAAGATCTATTGA 120
Qy 1284 AGGAGTTCGAATATATTTGTCAGAGCTGCGCTACAGATAGTGTGATCTCACTGAATATTAT 1343
Db 121 AGGAGTTCGAATATATTTGTCAGAGCTGCGCTACAGATAGTGTGATCTCACTGAATATTAT 180
Qy 1344 CCTACTAGATCGCTCCAGACTCGCCCTACAGATAGTGTGATCTCACTGAATATTAT 1403
Db 181 CCTACTAGATCGCTCCAGACTCGCCCTACAGATAGTGTGATCTCACTGAATATTAT 240
Qy 1404 CCCAGTGAAGATGATGTTATGGAAGCGCAGAGACTCATTTGAGTCACTGCTGATTTCTGT 1463
Db 241 CCCAGTGAAGATGATGTTATGGAAGCGCAGAGACTCATTTGAGTCACTGCTGATTTCTGT 300
Qy 1464 GAGGCACATTAATATCTATGAAGAAACAACAGACATCTGGATAAATATCCATGACATCTT 1523
Db 301 GAGGCACATTAATATCTATGAAGAAACAACAGACATCTGGATAAATATCCATGACATCTT 360
Qy 1524 TCATGTTTTTCCCAAGTCAAGAAAGGAAATTTGAGTTTATTTTGGCTCTGAAATGCAA 1583
Db 361 TCATGTTTTTCCCAAGTCAAGAAAGGAAATTTGAGTTTATTTTGGCTCTGAAATGCAA 420
Qy 1584 AACAGGTTTCCGTCAATTTATACAAATTAATCTATTTTAAAGAAAGCAATATTAACG 1643
Db 421 AACAGGTTTCCGTCAATTTATACAAATTAATCTATTTTAAAGAAAGCAATATTAACG 480
Qy 1644 ATCCAGTGTGGGCTGCTGCTCCAAAGTGAATTTCAAGTGTCTTATCAAGAGAGATAGC 1703
Db 481 ATCCAGTGTGGGCTGCTGCTCCAAAGTGAATTTCAAGTGTCTTATCAAGAGAGATAGC 540
Qy 1704 AATTACCAAGTGTGAATGGGAAGTTCTTGGCCGCGCATGGATCTAATATCAAGTTGATGA 1763
Db 541 AATTACCAAGTGTGAATGGGAAGTTCTTGGCCGCGCATGGATCTAATATCAAGTTGATGA 600
Qy 1764 AGTCAGAAGGCTGGTATATTTTGAAGGCACCAAGAGACTCCCTTTAGAGCATCACCTGTA 1823
Db 601 AGTCAGAAGGCTGGTATATTTTGAAGGCACCAAGAGACTCCCTTTAGAGCATCACCTGTA 660
Qy 1824 CGTAGTCAAGTACGTAAATCCTGGAGAGGTGACAAAGCTGACCGTGGCTACTCACA 1883
Db 661 CGTAGTCAAGTACGTAAATCCTGGAGAGGTGACAAAGCTGACCGTGGCTACTCACA 720
Qy 1884 TTCTGCTGCATCACTGACGACTGTGACTTCTTTATAGTAAGTATAGTAACCAAGAA 1943
Db 721 TTCTGCTGCATCACTGACGACTGTGACTTCTTTATAGTAAGTATAGTAACCAAGAA 780
Qy 1944 TCACACTGTGTGCTCTTTTACAGCTATCAAGTCTGGAAGATGACCCCACTTGCRAAAC 2003
Db 781 TCACACTGTGTGCTCTTTTACAGCTATCAAGTCTGGAAGATGACCCCACTTGCRAAAC 840
Qy 2004 AAAGGAATTTTGGGCCACCATTTTGGATTTCAGAGGCTCTCTTCTGCTACTACTCTCC 2063
Db 841 AAAGGAATTTTGGGCCACCATTTTGGATTTCAGAGGCTCTCTTCTGCTACTACTCTCC 900
Qy 2064 AGAAATTTTCTCTTTTGAAGTACTACTGGAATTTACATTTATGGAATGCTCTCAAGCC 2123
Db 901 AGAAATTTTCTCTTTTGAAGTACTACTGGAATTTACATTTATGGAATGCTCTCAAGCC 960
Qy 2124 TCATGATCTACAGCCTGGAAAGAAATATCCTACTGTGCTGTTTCATATATGTTGCTCTCA 2183
```


574	CCCAGTCACTCTCTGTGGATCTTATGATACAGGATACACGGAAGTTATATGGTCACC	633
2561	CTGACAGAAATGAACAGGGGCTATTAATTCTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCC	2620
634	CTGACAGAAATGAACAGGGGCTATTAATTCTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCC	693
2621	CCCTCTGAACCAAAATCGTTTACTGCTCTTATACATGGTTTCCCTGGATGAGAATCTCCATTTTG	2680
694	CCCTCTGAACCAAAATCGTTTACTGCTCTTATACATGGTTTCCCTGGATGAGAATCTCCATTTTG	753
2681	CACATACCAGTATATTACTAGATGTTTTTACTGAGGGCTGGAAGGCCATATGATTTACAGA	2740
754	CACATACCAGTATATTACTAGATGTTTTTACTGAGGGCTGGAAGGCCATATGATTTACAGA	813
2741	TCGTATCCTCAGGAGACACAGATAGAGTTCCTTGAATCGGAGAAATTAATCAACTGC	2800
814	TCGTATCCTCAGGAGACACAGATAGAGTTCCTTGAATCGGAGAAATTAATCAACTGC	873
2801	ATCTTTTGCATCACTTCCTCAGAAAACCTTGGATCAGTATTGCTCTCTTAAAGTGATAT	2860
874	ATCTTTTGCATCACTTCCTCAGAAAACCTTGGATCAGTATTGCTCTCTTAAAGTGATAT	933
2861	AAATTTTGACCTGTGTAGAACTCTCTGTATACATCTGGCTATTATTACCAAAATGAGAGGTT	2920
934	AAATTTTGACCTGTGTAGAACTCTCTGTATACATCTGGCTATTATTACCAAAATGAGAGGTT	993
2921	TAAACAACAGAAAACACAGAAATGATCATCATATTTTGATACCTGCCATGTAACATCTAC	2980
994	TAAACAACAGAAAACACAGAAATGATCATCATATTTTGATACCTGCCATGTAACATCTAC	1053
2981	TCTGAAAATAAATGTGTGTCATGCAAGGGTCTACGGTTTGTGTAGTAAATCTAATACC	3040
1054	TCTGAAAATAAATGTGTGTCATGCAAGGGTCTACGGTTTGTGTAGTAAATCTAATACC	1113
3041	TTAAACCCCAATGCTCAAAATCAAATGATATATTTCTGTAGAGAGCCACGAAATACCATA	3100
1114	TTAAACCCCAATGCTCAAAATCAAATGATATATTTCTGTAGAGAGCCACGAAATACCATA	1173
3101	AGAAATTACTATAAAAAAAAAA	3120
1174	AGAAATTACTATAAAAAAAAAA	1193

RESULT 11

```

US-09-976-674-10
; Sequence 10, Application US/09976674
; Patent No. 684180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 1356
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-976-674-10

```

Db	1	AAGTCTAAAGCTCTCGAGGCCAAGCGCGTCTACTGCGCGCGTCTTCTTAGTGGCG	60
Qy	61	CGTTTCGCGCGCTCGGTTGTTCACCGCGCGCGCGCGAGNAGCCACTCGCAACACGAGCG	120
Db	61	CGTTTCGCGCGCTCGGTTGTTCACCGCGCGCGCGCGAGNAGCCACTCGCAACACGAGCG	120
Qy	121	GAGTGGAGCGCGCGACGATGAAGCGCGCGAGCGCGCTCCATAGCGCACTGTGGGACGG	180
Db	121	GAGTGGAGCGCGCGACGATGAAGCGCGCGAGCGCGCTCCATAGCGCACTGTGGGACGG	180
Qy	181	TCGCGCGCGGCGCGGGGGAAGAAAAATGCAACATGCGCAGCAATGGAAAAACAGAACAG	240
Db	181	TCGCGCGCGGCGCGGGGGAAGAAAAATGCAACATGCGCAGCAATGGAAAAACAGAACAG	240
Qy	241	CTGGGTGTTGAGATATTTGAAACTGCGGACTGTGAGGAGAAATATTGAAATCAGAGATCGG	300
Db	241	CTGGGTGTTGAGATATTTGAAACTGCGGACTGTGAGGAGAAATATTGAAATCAGAGATCGG	300
Qy	301	CCCTAAATTCGGAGCCTTTTTATGTTGAGCGGTATTCTCTGGAGTCAGACTTAAAAAGCTGCTT	360
Db	301	CCCTAAATTCGGAGCCTTTTTATGTTGAGCGGTATTCTCTGGAGTCAGACTTAAAAAGCTGCTT	360
Qy	361	GCGGATACGAGAAATATCATGGCTACATGATGGCTAAGGCACCAATGATTTTCATGTTT	420
Db	361	GCGGATACGAGAAATATCATGGCTACATGATGGCTAAGGCACCAATGATTTTCATGTTT	420
Qy	421	GTGAGAGGAATGATTCAGATGGACCTCATTCAGACAGAAATCTATTACCTTGCCATGCT	480
Db	421	GTGAGAGGAATGATTCAGATGGACCTCATTCAGACAGAAATCTATTACCTTGCCATGCT	480
Qy	481	GGTGGAACACAGAAAAATACACTGTTTATTCTGAAATTCCTAAAACTATCAATAGAGCA	540
Db	481	GGTGGAACACAGAAAAATACACTGTTTATTCTGAAATTCCTAAAACTATCAATAGAGCA	540
Qy	541	GCAGTCTTAATGCTCTCTCTGGAAGCCTCTTTTGGATCTTTTCAGGCAACACTGGACTAT	600
Db	541	GCAGTCTTAATGCTCTCTCTGGAAGCCTCTTTTGGATCTTTTCAGGCAACACTGGACTAT	600
Qy	601	GGAACTATTCTCGAGAGAGAAAGAACTATTAGAGAAAAAGAAACGCAATTTGAAACAGTCGGA	660
Db	601	GGAACTATTCTCGAGAGAGAAAGAACTATTAGAGAAAAAGAAACGCAATTTGAAACAGTCGGA	660
Qy	661	ATTGCTCTTTAGATATATCACAGGAGTGGACATTTCTGTTTCAAGCCGGTAGTGGGA	720
Db	661	ATTGCTCTTTAGATATATCACAGGAGTGGACATTTCTGTTTCAAGCCGGTAGTGGGA	720
Qy	721	ATTTATCAGTAAAAAGATGGAGGCCACAGGANTTTACGCAACAACTTTTAAGGCCCAAT	780
Db	721	ATTTATCAGTAAAAAGATGGAGGCCACAGGANTTTACGCAACAACTTTTAAGGCCCAAT	780
Qy	781	CTAGTGGAAACTAGTTGTGCCAACATACGGAATGGATCCAAAATTTATGCCCGCTGATCCA	840
Db	781	CTAGTGGAAACTAGTTGTGCCAACATACGGAATGGATCCAAAATTTATGCCCGCTGATCCA	840
Qy	841	GACTGGATTGCTTTTATACATAGCAACAGATATTTGGATATCTAACATCGTAAACAGAGAA	900
Db	841	GACTGGATTGCTTTTATACATAGCAACAGATATTTGGATATCTAACATCGTAAACAGAGAA	900
Qy	901	GAAGAGAGACTCACTTATGTGCACAATG	928
Db	901	GAAGAGAGACTCACTTATGTGCACAATG	928

RESULT 12

```

RESOLUT 12
US-09-976-674-4
; Sequence 4, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPP1V

```


FILE REFERENCE: 70669
CURRENT APPLICATION NUMBER: US/09/976.674
CURRENT FILING DATE: 2001-10-12
PRIOR APPLICATION NUMBER: US 60/240,117
PRIOR FILING DATE: 2000-10-12
NUMBER OF SEQ ID NOS: 61
SOFTWARE: PatentIn version 3.1
SEQ ID NO 4
LENGTH: 2617
TYPE: DNA
ORGANISM: Homo sapiens
US-09-976-674-4

Query Match 27.6%; Score 861; DB 3; Length 2617;

Best Local Similarity 59.4%; Pred. No. 6.9e-420;
Matches 1499; Conservative 0; Mismatches 1420; Indels 6; Gaps 2;

Qy	316	TTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGAGCTGCTCCGATACCAAGAAA	375
Db	80	TTCCAGGTGCGAAGCACTCGTGGACGGGCTCCGGAGCATCATCCACGGCGCCGCAAG	139
Qy	376	TATCATGCTCATGATGGCTTAAGGCACCATGATTTTCATGTTTGTGAAGAGGAATGAT	435
Db	140	TACTCGGCGCTCATGTCAACAGCGCCCAAGCTTTCAGTTTGTGACAGAGCGGAT	199
Qy	436	CCAGATGACCTCATTTACAGACAGAACTTATTAACCTTGCCATGTCTGGTGAGAACAGAAA	495
Db	200	GAGTCTGGGCGCCACTCCACCGGCTCTACTACCTGGGAATGCCATATGGCAGCCGAGAG	259
Qy	496	AATACACTGTTTATTTCTGAAATTTCCAAAATCATCAATAGACAGCAGCTTTAAATGCTC	555
Db	260	AATCTCCCTCTACTCTGAGATTTCCCAAGAGTCCGGAAGAGGCTCTGCTGCTCCTG	319
Qy	556	TCTTGAAGAGCTCTTTTGGATCTTTTTCAGGCAACACTGGAATGATGATTTCTGA	615
Db	320	TCCTGGAAGCAGATGCTGGATCATTTCCAGGCGCACGCCCAACCATGGGGTCTACTCTCG	379
Qy	616	GAAGAGAACTTAAAGAAAGAAAGCAATTTGGACAGCTCGGAATTCCTTTTACGAT	675
Db	380	GAGGAGGAGCTGCTGAGGAGCGGAAACGCTCGGGGTCTTGGCATCATCCTCTCAAGAC	439
Qy	676	TATCAACCAAGGAGTGAACATTTCTGTTTCAAGCGGTAGTGGAAATTTATCAAGTAAA	735
Db	440	TTCCACAGCAGAGTGGCTCTTCTCTTCCAGGCGAGCAAGACCTCTTCCACTGTGCG	499
Qy	736	GATGAGGGCCACAGGATTTTACCAACAACTTTTAAGGCCCAATCTAGTGGAAACTAGT	795
Db	500	GACGGCGCAAGACCGCTTCATGGTGTCCCTATGAAACCGCTGGAAATCAAGACCCAG	559
Qy	796	TGTCCTCAACATACGAGTGGATCCAAAATTTATGCCCGCTGATCCAGACTGGATGCTTTT	855
Db	560	TGCTCAGGGCCCGGATGGACCCCAATCTGCGCTCCGACCCCTGCTTCTTCTCTTC	619
Qy	856	ATACATACAAACGATTTTGGATATCTAAATCTGTAAACAGAGAAAGAGACTCACT	915
Db	620	ATCAATAACAGCAGCTGTGGTGGCCAAATCATGAGACAGGCGGAGGAGCGGGCTGACC	679
Qy	916	TATGTGCAATGAGCTAGCCAACTAGGAAGAGATGCCAGATCAGCTGGAGTCGCTACC	975
Db	680	TTCTGCCCAAGGTTTATCAATGTCTCTGGATGACCCCAAGTCTCGGGGTGGGCCACC	739
Qy	976	TTTGTCTTCAAGAGAAATTTGATGATATTTCTGGCTATTGGTGGTGTCCAAAGCTGAA	1035
Db	740	TTCTGTATACAGGAAGAGTTCGACCGCTTCACTGGGTACTGGGTGGTGGCCCAAGCTCC	799
Qy	1036	AC---AATCCCGAGTGGTGAATAATTTCTTAGAATTTCTATGAGAAATGATGATCT	1092
Db	800	TGGGAAGTTTACAGAGGCGCTCAAGACGCTGCGAATCTCTGTATGAGGAATGCGATGCTCC	859
Qy	1093	GAGGTGGAAATTTATGTTTACATCTCCCTATGTTGGAAACAAAGAGGGGAGATTTCAATC	1152
Db	860	GAGGTGGAGGTCAATTCAGTCTCCCTCTCTCTGGCTAGAGAAAGGAAGAGCGGACTGAT	919

Qy	1153	CGTTATCTTAAACAGGTACAGCAAAATCCTAAAGTCACTTTTAAAGATGTGAGAAATATG	1212
Db	920	CGGTACCCAGACAGGAGCAAGAAATCCCAAGATTCCTTTGAAATGGCTGAGTTCAG	979
Qy	1213	ATTGATGCTGAAGGAAGATCATAGATGTCTAGATAGGAACATAATTCAACTTTTGGAG	1272
Db	980	ACTGACAGCCAGGGCAAGATCGTCTCGACCCAGGAGAGGAGCTGGTGCAGCCCTTCAGC	1039
Qy	1273	ATTCTATTGAAGGAGTTGAATATATTCAGAGAGCTGGATGAGTCTCTGAGGGGAAAATAT	1332
Db	1040	TGCTCTGTTCCCGAAGGTGGAGTACATCCGACGGGCGGGTGGACCCGGGATGGCAATATC	1099
Qy	1333	GCTTGGTCCATCTACTAGATCGCTCCAGACTCGCTTACAGATAGTGTGATCTCACT	1392
Db	1100	GCTGGGCGATGTTCTTGACCGGCGCCAGCAGTGGCTCCAGCTCGTCTCTCCCGCG	1159
Qy	1393	GAATTAATTTATCCAGTGAAGATGATTTATGAAAGGCGAGAGACTCATTTGAGTCAAGT	1452
Db	1160	GCCCTGTTTCATCCCGACACAGAAATGAGGAGCGGGCTAGCTCTGCCAGAGCTGTC	1219
Qy	1453	CTGATTTGTGACGCGCACTAATTAATCTATGAAGAAACACAGACATCTGGATTAATATC	1512
Db	1220	CCAGGAATGTCCAGCGGTATGTGTGTACGAGGAGGTCAACAACTGTGGATCAATGTT	1279
Qy	1513	CATGACATCTTTCATGTTTTCCTCCAAAGTCAAGAGAG---GAAATGAGTTTTATTTT	1569
Db	1280	CATGACATCTTTCATGTTTTCCTCCCAATCAGAGGAGAGGAGCTCTGCTTTTCTCCGC	1339
Qy	1570	GCCTCTGAATGAAACAGGTTTCGTCATTTATACAAAATTAATACATCTATTTAAAGGAA	1629
Db	1340	GCCAAATGAATGCAAGACCGGCTTCTGCCATTTGTACAAAGTCAACGCGGTTTTAAATCC	1399
Qy	1630	AGCAATATAAAGCATCCAGTGGTGGCTGCTGCCAAGTGTATTTCAAGTGTCTTATC	1689
Db	1400	CAGGGCTACGATTTGAGTGGAGCGCTTCAGCCCGGGGAGAGATGAATTTAGTGGCCCAT	1459
Qy	1690	AAAGAGGAGATAGCAATTTACAGTGGTGAATGGGAAGTTTCTGGCCGCAATGATCTAAT	1749
Db	1460	AAGGAGAGATTTGCTCTGACCGGTGAATGGGAGGTTTGGCGAGGACCGCTCCAG	1519
Qy	1750	ATCAAGTTGATGAAGTCAAGAGGTGGTATATTTTGAAGGACCAAGAGATCCCTTTTA	1809
Db	1520	ATCTGGGTCAATGAGGAGACCAAGCTGTGTACTTCCAGGCGACCAAGGACACGCGCTG	1579
Qy	1810	GAGCATCACTGATCGTAGTCAATTAATCTGAGAGAGTGAACAGGCTCACTGAC	1869
Db	1580	GAGCACACCTCTACGTGTCTAGCTATGAGGCGGCGCGGAGATCGTACGCTCACCACG	1639
Qy	1870	CGTGGCTACTCATTTCTTGTGCTGATCAGTCAAGCTATCAAGCTATCAAGTAAAGTAT	1929
Db	1640	CCCGGCTCTTCCCATAGCTGCTCCATGAGCCAGAACTTCCGACATGTTGCTCAGCCACTAC	1699
Qy	1930	AGTAAACAGAGAAATCCACACTGTGTCTCTTTTACAAAGTATCAAGTCTCTGAAGATGAC	1989
Db	1700	AGCAGGCTGAGCACGCGCCCTTGGTGCACGCTCTCAAGCTGAGGGGCGCCGACGAC	1759
Qy	1990	CCAATTTGCAAAACAAAGGAAATTTTGGGCCCACTTTTGGATTTGAGAGGTCCTCTTCT	2049
Db	1760	CCCTGCAACAGCAGCCCGCTTCTGGGCTAGCATGATGAGGCGAGCCAGCTGCCCCCG	1819
Qy	2050	GACTATCTCTCCAGAAATTTTCTCTTTGAAAGTACTACTGGAATTTACATGTTATGGG	2109
Db	1820	GAATAGTTCTCCAGAGATCTTCCATTTCCACACGCGCTCGGATGTGCGGCTCTACGCG	1879
Qy	2110	ATGCTCTACAGGCTCATGATCTACAGCTGGAAGAAATATCTCTATGTTGCTGTTCTATA	2169
Db	1880	ATGATCTACAGGCCCCAGCCTTGCAGCCAGGGAAGAGCACCCACCGCTCTCTTTGTA	1939
Qy	2170	TATGTTGGTCTCAGGTGAGTGTGTAATTAATCGTTTAAAGGAGTCAAGTATTTCCGC	2229
Db	1940	TATGAGGGCCCCAGGTGAGTGTGTAATTAATCTCTTCAAGGCAATCAAGTACTTGGCG	1999
Qy	2230	TTGAACTACCTAGCCTCTCTAGGTTATGTTGGTTGTAGTATAGACAAACAGGGGATCTCTG	2289

Db 2000 CTCACACACTGGCTCCCTGGGCTAGCGCGTGTGTGATTTGACGGCAGGGGCTCCCTGT 2059
Qy 2290 CACGAGGGCTAAATTTGAGGGGCTTTAAATATATAATATGAGTCAATAGAAATTCAC 2349
Db 2060 CAGGAGGGCTTCGGTTGAGAGGGGCTCTGAAACCAATATGGCCAGGTGGAGATCGAG 2119
Qy 2350 GATCAGGTGGAAGACTCCAAATATCTAGCTTCTCGATATGATTTTCAATGACTAGATGGT 2409
Db 2120 GACAGGTGGAAGGGCTTCGAGTTCGTGGCCGAGAGATATGGCTTCATCGACTGAGCCGA 2179
Qy 2410 GTGGGATCCACGGCTGGTCTATGAGAGGATACCTCTCCCTGATGGCATTAATGAGAGG 2469
Db 2180 GTTGCCATCCATGGCTGGTCTACGGGGGCTTCTCTCGCTCATGGGGCTAATCCACAAG 2239
Qy 2470 TCAGATATCTTCAGGGTGTCTATGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGAT 2529
Db 2240 CCCAGGTGTTCAAGGTGGCCATCGCGGGTGGCCCCGGTCAACGCTCTGGATGGCTACGAC 2299
Qy 2530 ACAGGATACAGGGAACGTATATGGGTCACTCTGACAGAAAGAAAGAGGCTATTACTTA 2589
Db 2300 ACAGGGTACATGAGCGCTACATGAGCGTCCCTGAGAACAAACAGCACGGCTATGAGCG 2359
Qy 2590 GGATCTGTGGCCATGCAAGCAGAAAGTTCCTCTGAAACCAATCGTTTACTGCTCTTA 2649
Db 2360 GGTTCGTGGCTTCGACGTGGAGAGCTGCCCAATGAGCCCAACCGCTTGTATCTCTC 2419
Qy 2650 CATGTTCTCCGATGAGATGCTCATTTTGCATACCATGATATATGAGTGTTTTA 2709
Db 2420 CACGGCTTCCTGGAGAAACGTGACCTTTTTCACACAACTTCTCTGCTCCCACTG 2479
Qy 2710 GTGAGGGTGAAGAGCCATATGATTTACAGATCTATCTCCAGGAGAGACACAGATAAGA 2769
Db 2480 ATCCGAGCAGGGAACCTTACAGCTCCAGATCTACCCCAAGAGAGACACAGTATTCG 2539
Qy 2770 GTTCTGTAATCGGAGAACATATGAACTGCATCTTTTGCACACTTCAAGAAACCTT 2829
Db 2540 TGCCCCGATCGGGCGAGCACTATGAAGTCACGTGTGCTGACACTTTCTACAGGAATACCTC 2599
Qy 2830 GGATC 2834
Db 2600 TGAGC 2604

RESULT 13

; Sequence 28, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsaya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 28
; LENGTH: 4219
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-976-674-28

Query Match 27.6%; Score 861; DB 3; Length 4219;

Best Local Similarity 59.4%; Pred. No. 9.3e-242;
Matches 1499; Conservative 0; Mismatches 1020; Indels 6; Gaps 2;

Qy 316 TTTTATGTGAGCGGTATTCCTGGAGTCAGCTTAAAGCTGCTTCGGATACGAGAAA 375

Db 436 TTCCAGGTGCAGAAAGCACTCCTGGGACGGGCTCCGAGCATCATCCACGGCAGCCGCAAG 495
Qy 376 TATCATGGCTACATGATGGCTAAGGCACCAACATGATTTTCTTTGTAAGAGGAATGAT 435
Db 496 TACTCGGGCTCATTTGTCAACAGAGCGCCCGACGACTTCCAGTTTGTGCAAGAGCGGAT 555
Qy 436 CCAGATGGACCTCATTTACAGACAGAAATCTATTACCTTCGCATGTCTGGTGAGAACAGAGAA 495
Db 556 GAGTCTGGGGCCCACTCCACCGCTCTACTACCTGGGAATGCCATATGCGACCGGAGAG 615
Qy 496 AATACTACTGTTTATTTCTGAAATTTCCCAAACTATCAATAGACGACAGCTCTTAATGCTC 555
Db 616 AACTCCCTCTCTACTCTGAGATTTCCCAAGAGGTCCGGAAGAGGCTCTGCTGCTCTG 675
Qy 556 TCTTGAAAGCTCTTTTGGATCTTTTTCAGGCAACTGGACTATGAAATGTAATCTCGA 615
Db 676 TCTTGAAAGCAGATCTGGATCATTTTCAGGCCACGCCACCACTGGGGTCTACTCTCG 735
Qy 616 GAAGAAGAACTATTAAAGAGAAAGAAAGCAAGTTCGAAACAGTTCGGAATTTGCTTCTTACGAT 675
Db 736 GAGGAGGAGCTGCTGAGGGAGCGGAAAGCGCTGGGGTCTTCGGCATCACCTCTCTACGAC 795
Qy 676 TATCAACAGGAAGTGGAAACATTTCTGTTTCAAGCCGCTAGTGGAAATTTATCAGCTAAAA 735
Db 796 TTCCAAGCAGAGTGGCTCTTCTCTTCCAGGCCAGCAACAGCTCTTCCACTGCGC 855
Qy 736 GATGAGGGCCACAAAGGATTTACGCAACAACTTTAAAGGCCCAATCTAGTGGAAACTAGT 795
Db 856 GACGGCGCAAGAACGGCTTCATGGTGTCCCTATGAACCGCTGGAAATCAAGACCCAG 915
Qy 796 TGTCCCAACATACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 855
Db 916 TGTCTAGGGCCCGGATGGACCCCAAAATCTGCCCTGCGGACCTGCTCTTCTCTCTC 975
Qy 856 ATACATAGCAAGATTTTGGATCTTACATCTTACATCTTACATCTTACATCTTACATCTTACAT 915
Db 976 ATCAATACAGCGACTGTGGTGGCCAACTCAGACAGCGAGAGCGCGGCTGACC 1035
Qy 916 TATGTGCAATGAGCTAGCCAAACATGGAAGAGATGCCAGATCAGCTGAGTGGCTACC 975
Db 1036 TTTGCGCACCAAGGTTTATCCAATGTCTGATGACCCCAAGTCTGCGGGTGGGCCACC 1095
Qy 976 TTTGTTCTCCAAAGAAATTTGATAGATATTTCTGGCTATTGGTGGTGTCCAAAGCTGAA 1035
Db 1096 TTTGCTCATACAGAAAGAGTTCCACCGCTTCACTGGGTA CTGGTGGTGGCCCAACAGCTCC 1155
Qy 1036 AC---AATCTCCAGTGGTGGTAAATTTCTTAAAGTTCTTATATGAAGAAATGTAATCT 1092
Db 1156 TGGGAAGGTTACAGAGGGCTCAAGACGCTGCGAAATCTGTATGAGGAAGTCGATGAGTCC 1215
Qy 1093 GAGGTGGAATTTATTCATGTTTACATCCCTATGTTGGAACAAGAGGGGAGGAGATTCATTTC 1152
Db 1216 GAGTGGAGGTCAATTCACGTCCCTCTCTCGCTTAGAAGAAAGAAAGACGGACTCGTAT 1275
Qy 1153 CGTTATCTTAAACAGGTACAGCAATCTTAAAGTCACTTTTAAAGATGTGAGAAATATG 1212
Db 1276 CGGTACCCAGGACAGGAGCAAGANTCCCAAGATTCCTTGAATGGCTGAGTTCCAG 1335
Qy 1213 ATTGATGTGAAGAGGATCATAGATGTATAGATAGGACATTAATTCACCTTTTGG 1272
Db 1336 ACTGACAGCCAGGCAAGATCGTCTCGACCCAGGAAGAGCTGTTGCGAGCCCTTCAGC 1395
Qy 1273 ATTCATTTGAAGAGGTGAATATATATCCAGAGCTGGATGGACTCTCTGAGGAAATAT 1332
Db 1396 TCGCTGTTCCGAAGGTGGAGTACATCGCAGGGCCGGTGGACCCGGGATGGCAATAC 1455
Qy 1333 GCTTGTCCATCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGTGATCTCACT 1392
Db 1456 GCCTGGGCAATGTTCTGGACCGGCCCGAGAGTGGCTCCAGCTGCTCTCTCTCCCCCG 1515
Qy 1393 GAATTTATTTATCCAGTGAAGATGATGTTATGGAAGGCGAGAGACTCATTTAGTCAAGT 1452
Db 1516 GCCTGTTTATCCGAGCAGAGAAATGAGGAGCAGCGGCTAGCTCTCTGCGAGAGCTGTC 1575

Db 736 GAGGAGAGCTGCTGAGGAGCGGAAACGCTGGGGTCTTCGGCATCACTCTCAAGC 795
Qy 676 TATCA CCAAGGAAGTGAA CATTTCTGTTTCAAGCGGTAGTGGAAATTTATCAGTAA 735
Db 796 TTCCACAGCGAGAGTGGCTTCTCTCTCCAGGCGAGCAACAGGCTCTTCCACTGCCG 855
Qy 736 GATGAGGCGCAACAGGATTTACGCAACAACTTTTAAGGCCCAATCTAGTGGAAACTAGT 795
Db 856 GACGCGCGCAAGAACGCTTTCATGGTGTCCCTTATGAACCGCTGGAAATCAAGACCCAG 915
Qy 796 TGTCCCAACATACGGATGGATCCAAATATGCCCCGCTGATCCAGACTGGATGCTTTT 855
Db 916 TGCTCAGGCGCCGATGGAGCCCAAAATCTGCCCTGCCGACCTGCTCTCTCTCTCTC 975
Qy 856 ATACATAGCAACAGATTTTGGATATCTAAATCGTTAAACGAGAGAAAGAGACTCACT 915
Db 976 ATCAATACAGGACCTGTGGTGGCCCAACATCGACAGGCGAGGCGGCTGACC 1035
Qy 916 TATGTGCAATGAGCTAGCCACATCGAAGAGATGCCAGATCAGCTGAGTGCCTACC 975
Db 1036 TTCTGCCACCAAGTTTATCCATGTCTCGATGACCCCAAGTCTCGGGTGTGGCCACC 1095
Qy 976 TTGTGTTCTCAAGAGAAATTTGATAGATATCTGGCTATTTGGGTATGGTGTCCAAAGCTGAA 1035
Db 1096 TTCTGTCATACAGGAGAGTTCGACCGCTTCACTGGGTACTGGTGGTGGCCCAAGCTCC 1155
Qy 1036 AC--AACTCCAGTGGGTGTAATAATCTTAGAATTTCTATATGAAGAAATGATGAATCT 1092
Db 1156 TGGGAAGTTTCAAGAGGCGCTCAAGACGCTCGGAATCTGTATGAGGAGTCGATGAGTCC 1215
Qy 1093 GAGGTGGAATTTATCTATGTTTACATCCCTTATGTTGGAACAAAGAGGCGAGATTCATTC 1152
Db 1216 GAGGTGGAGGTCAATTCAGCTCCCTCTCTCGCTAGAGAAAGGAAGACGAGCTCGTAT 1275
Qy 1153 CGTTATCTCTAAACAGGTACAGCAAAATCTCTAAAGTCTCATTTTAAGATGTGAGAAATATG 1212
Db 1276 CGGTATCCCAAGCAGGAGCAAGAAATCCCAAGATTTGCTTGAATCGCTGAGTTCAG 1335
Qy 1213 ATTGATGCTGAAGGAAGATCATAGATGTATAGATAAGGAATTAATTCACCTTTTGA 1272
Db 1336 ACTGACAGCGGCGAAGATCGTCTCGACCCAGGAGAGGAGCTGGTGGAGGCTTCAGC 1395
Qy 1273 ATTCTATTGAAGGATTTGAATATATTTGCCAGAGCTGGATGGACTCTCAGGGAATAAT 1332
Db 1396 TCGCTGTTCCGNAAGTGGATACATCGCAGGCGGGTGGACCCGGGATGGCAATAC 1455
Qy 1333 GCTTGGTCCATCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACT 1392
Db 1456 GCTGGCCATGTTCTTGACCGGCCCCAGCAGTGGCTCAGCTCGTCTCTCTCCCCCG 1515
Qy 1393 GAATTAATTTATCCAGTGAAGATGATGTTATGGAAGGCGAGACTCATGAGTCAGTG 1452
Db 1516 GGCCTGTTATCCGAGCACAGAGAAATGAGGAGCGGGCTAGCCTCTGCGAGCTGTC 1575
Qy 1453 CTGATTTCTGTGACGCCACTAATTTATGAGAGAAACACAGACATCTGGATTAATATC 1512
Db 1576 CCAGGAATGTCCAGCCGATGTGGTGTACAGAGAGTCCAAACGCTGGATCAATGTT 1635
Qy 1513 CATGACATCTTTTATGTTTTTCCCAAGTCCAGAG--GAAATGAGTTTATTTTT 1569
Db 1636 CATGACATCTTTCTATCTCTCCCAATCAGAGGAGAGGAGCTGCTGTTCTCCGC 1695
Qy 1570 GCTCTGAAATGCAAAACAGGTTTCGTCATTTTACAAATTAATCATCTATTTTAAAGGAA 1629
Db 1696 GCCAATGAATCAAGACCGGCTTCTGCCATTTGTACAAAGTCAACCGCGTTTAAATCC 1755
Qy 1630 AGCAATATAAAGCATTCAGTGGTGGCTGCTGCTCAAGATGATTTCAAGTGTCTATC 1689
Db 1756 CAGGGCTACGATTTGAGTGAGCCCTTCAGCCCGGGGAAGATGAATTAAGTCCCAAT 1815
Qy 1690 AAAGAGGAGATAGCAATACAGTGGTGAATGGGAAGTTCTTGCGCGGATGGATCTAAT 1749

Db 1816 RAGGAAGAGATTGCTCTGACCCAGCGGTGAATGGAGGTTTGGCGAGGACGCTCCAAG 1875
Qy 1750 ATCCAAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGACCAAGACTCCCTTTTA 1809
Db 1876 ATCTGGGTCAATGAGGAGACCAAGCTGGTGTACTTCCAGGCGACCAAGGACACGCGCTG 1935
Qy 1810 GAGCATCACCTGTAGTACGTAAATCCCTGGAGAGGTGACAAAGGCTGACTGAC 1869
Db 1936 GAGCACCACTCTACGTGGTCACTATGAGCGCGCGCGGATCGTACGCTCAACGAC 1995
Qy 1870 COTGCTACTCATCTTCTGCTGATCAGTCAGCTGATCTTCTTTAAGTAGATAT 1929
Db 1996 CCGGCTTCTCCATAGCTCTCCATGAGCGAGACTTCGACATGTTCTGACGCCACTAC 2055
Qy 1930 AGTAAACCAAGAAATCCACACTGTGTCTCTTTTCAAGCTATCAAGTCTGAGATGAC 1989
Db 2056 AGCAGCTGAGCAGCGCGCTCTGCTGCACTGTACAGCTGAGCGCGCCGACGAC 2115
Qy 1990 CCAACTTGCAGAAACAAAGGAAATTTTGGGCCACCAATTTTGGATTCAGCAGGTCTCTTC 2049
Db 2116 CCGCTGCAACAGCAGCCGCTTCTGGGCTAGCATGATGAGGCGAGCAGCTGCCCCCG 2175
Qy 2050 GACTATCTCTCCAGAAATTTCTCTTTTGAAGTACTACTGGATTTTACATTTGATGG 2109
Db 2176 GATTAATGTTCTCCAGAGATCTTCCATTTCCACACGCGCTCGATGTGGGCTCTACGCG 2235
Qy 2110 ATGCTCTCAAGCTCATGATCTACAGCTGGAAGAAATATCTCTACTGTGCTGTTTATA 2169
Db 2236 ATGATCTACAGCCCGCTTTCAGCGAGGAGAAAGCACCACCGTCTCTTTGTA 2295
Qy 2170 TATGCTGCTCAGCTGAGTGTGCTGTAATCGGTTTAAAGGAGTCAAGTATTTCCCG 2229
Db 2296 TATGAGGCGCCCGAGTGCAGCTGTGATTAATCTCTTCAAGGCATCAAGTACTTTCGG 2355
Qy 2230 TTGAATACCTTAGCTCTTAGTGTATGTTGTTAGTATAGACACAGGAGTCTCTGT 2289
Db 2356 CTCACACACTGGCTCTCTGGCTACCGCGTGTGTGATGACGCGAGGCTCTCTGT 2415
Qy 2290 CACGAGGCTTAAATTTTGAAGGCGCTTTAAATATAAATGAGTCAATAGAAATGAC 2349
Db 2416 CAGCAGGCTTCCGTTGGAAGGCGCTTGAAGGCGCTTGAAGGCGAGTGGAGATCGAG 2475
Qy 2350 GATCAGGTGGAAGGACTCAATATCTAGCTTCTCGATAGATTTTCAATTTGATAGTCTGT 2409
Db 2476 GACCAGGTGGAGGCTCGAGTTCTGGCGAGAGATATGGCTTCATCGACCTGAGCCGA 2535
Qy 2410 GTGGGATCCAGGCTGTCTCTATGGAGGATACTCTCCCTGATGGCATTTAATGCGAGG 2469
Db 2536 GTTCCCATCCATGCTGTCTTACGCGGCTTCTCTCGCTCATGGGCTTAATCCACAG 2595
Qy 2470 TCAGATATCTTTCAGGTTGCTATTTGCTGGGCGCCAGTCACTCTGTGGATCTTCTATGAT 2529
Db 2596 CCGCAGGTCTTCAAGGTGGCCATCGCGGCTGCCCGGCTCAACGCTCTGGATGGCTACGAC 2655
Qy 2530 ACAGGATACAGGAAGTTATATGGTCCACCTGACAGAAATGAACAGGCGTATTACTTA 2589
Db 2656 ACAGGATACCTGAGGCTACATGAGCGTCTCTGAGAACACAGCAGCTATGAGCG 2715
Qy 2590 GATCTGTGGCTACGAGCAAGGTTCCCTCTGACCAAAATCTGTTACTGCTCTTA 2649
Db 2716 GGTTCGTGGCTCTGACGTGGAGAGCTGCGCAATGAGCCCAACCGCTTCTTATCCTC 2775
Qy 2650 CATGTTTCTCGATGAGAAATGTCATTTTGGACATACCACTATTTACTGAGTTTTTA 2709
Db 2776 CAGGCTTCTGAGCAAGAACGTGCACTTTTCCACAACTTCTCTCGTCCCACTG 2835
Qy 2710 GTGAGGCTGGAAGCATAATGATTTTACATCTATCTCTGAGGAGACACAGCATAAGA 2769
Db 2836 ATCCGAGCAGGGAACCTTACAGCTCCAGATCTACCCCAACAGAGACACAGTATTCG 2895
Qy 2770 GTTCTGAAATCGGAGAACATTTATGAACTGTGACCTTTTGGACCTCTCAAGAAACCTT 2829
Db 2896 TGCCCGAGTCCGCGGAGCACTATGAAGTCAAGTGTGCTGCTGCTTCTACAGGAATACCTC 2955

Qy 2830 GGATC 2834
Db 2956 TGAGC 2960

RESULT 15

US-976-674-18
; Sequence 18, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akineanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
; LENGTH: 832
; TYPE: DNA
; ORGANISM: Homo sapiens
US-976-674-18

Query Match 25.7%; Score 801.4; DB 3; Length 832;
Best Local Similarity 99.5%; Pred. No. 1.le-224;
Matches 799; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGTGTAAAGCTCGAGGCCAAGCCGCTGCTACTGCCGCGCTCTTCTTAGTCCG 60
Db 1 AAGTGTAAAGCTCGAGGCCAAGCCGCTGCTACTGCCGCGCTCTTCTTAGTCCG 60

Qy 61 CGTTCCGCCCTGGGTGTCTACCGCGCGCGCGGAGGAGCCACTGCAACCCAGGACCG 120
Db 61 CGTTCCGCCCTGGGTGTCTACCGCGCGCGCGGAGGAGCCACTGCAACCCAGGACCG 120

Qy 121 GAGTGAGCGCGCAGCATGAAGCGCGCGCGCGCTCCATAGCGCAGCTCGGACCG 180
Db 121 GAGTGAGCGCGCAGCATGAAGCGCGCGCGCGCTCCATAGCGCAGCTCGGACCG 180

Qy 181 TCCGGCGCGCGCGCGGAGGAAATGCAACATGCGCAGCAGCAATGGAACAGAACAG 240
Db 181 TCCGGCGCGCGCGCGGAGGAAATGCAACATGCGCAGCAGCAATGGAACAGAACAG 240

Qy 241 CTGGGTGTTGAGATATTTGAAACTCGCGACTGTGAGGAGAAATTTGAATCACAGGATCGG 300
Db 241 CTGGGTGTTGAGATATTTGAAACTCGCGACTGTGAGGAGAAATTTGAATCACAGGATCGG 300

Qy 301 CCTAAATGGAGCCCTTTTATGTTGAGCGGTATTCCTGGAGTCACTTAAAGCTGCTT 360
Db 301 CCTAAATGGAGCCCTTTTATGTTGAGCGGTATTCCTGGAGTCACTTAAAGCTGCTT 360

Qy 361 GCCGATACAGAAATATCATGCTACATGATGCTAAGGCCACCATGATTTCAATGTTT 420
Db 361 GCCGATACAGAAATATCATGCTACATGATGCTAAGGCCACCATGATTTCAATGTTT 420

Qy 421 GTGAAGAGAAATATCATGCTACATGATGCTAAGGCCACCATGATTTCAATGTTT 480
Db 421 GTGAAGAGAAATATCATGCTACATGATGCTAAGGCCACCATGATTTCAATGTTT 480

Qy 481 GGTGAGACAGAGAAATACACTGTTTTTATCTGAAATTCCTAAATCTATCAATAGAGCA 540
Db 481 GGTGAGACAGAGAAATACACTGTTTTTATCTGAAATTCCTAAATCTATCAATAGAGCA 540

Qy 541 GCAGTCTTAATGCTCTCTTTGAGGCTCTTTTGGATCTTTTGGGCAACACTGGACTAT 600
Db 541 GCAGTCTTAATGCTCTCTTTGAGGCTCTTTTGGATCTTTTGGGCAACACTGGACTAT 600

Qy 601 GGAATGTATTTCTCGAGAAAGAACTATTAAAGAGAAAGAAACCGCATTTGGAACAGTCGGA 660
Db 601 GGAATGTATTTCTCGAGAAAGAACTATTAAAGAGAAAGAAACCGCATTTGGAACAGTCGGA 660

Qy 661 ATTGCTTCTTACGATTATCAACAGGAGTGGACATTTCTGTTTCAAGCGGTAGTGA 720
Db 661 ATTGCTTCTTACGATTATCAACAGGAGTGGACATTTCTGTTTCAAGCGGTAGTGA 720

Qy 721 ATTTATCAGTAAAGATGAGGGCCACAAGGATTTAGCGAAACAACTTTAAGGCCCAAT 780
Db 721 ATTTATCAGTAAAGATGAGGGCCACAAGGATTTAGCGAAACAACTTTAAGGCCCAAT 780

Qy 781 CTAGTGGAAACTAGTTGTCCCA 803
Db 781 CTAGTGGAAACTAGTTGTSCCA 803

Search completed: April 15, 2006, 00:11:57
Job time : 542 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 14, 2006, 22:13:06 ; Search time 2408 Seconds
(without alignments)
10714.476 Million cell updates/sec

Title: US-10-825-632-2
Perfect score: 3120
Sequence: 1 agtgcgttaagctccgagg.....agaattactataaaaaaaa 3120

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA Main:
1: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq:
3: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq:
4: /cgn2_6/prodata/1/pubpna/US09B_PUBCOMB.seq:
5: /cgn2_6/prodata/1/pubpna/US10A_PUBCOMB.seq:
6: /cgn2_6/prodata/1/pubpna/US10B_PUBCOMB.seq:
7: /cgn2_6/prodata/1/pubpna/US10C_PUBCOMB.seq:
8: /cgn2_6/prodata/1/pubpna/US10D_PUBCOMB.seq:
9: /cgn2_6/prodata/1/pubpna/US10E_PUBCOMB.seq:
10: /cgn2_6/prodata/1/pubpna/US11_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	3120	100.0	3120	7	US-10-415-122-5
2	3120	100.0	3120	8	US-10-825-632-2
3	3106.4	99.6	4829	3	US-09-976-674-12
4	3106.4	99.6	4829	3	US-10-982-512-12
5	3105.2	99.5	3143	6	US-10-170-789-37
6	3101.8	99.4	3106	7	US-10-311-035-30
7	2824.4	90.5	4685	3	US-09-976-674-22
8	2824.4	90.5	4685	3	US-10-982-512-22
9	2806.4	89.9	4676	3	US-09-976-674-20
10	2806.4	89.9	4676	3	US-10-982-512-20
11	2645.8	84.8	2649	6	US-10-170-789-39
12	2644.4	84.8	2671	3	US-09-976-674-2
13	2644.4	84.8	2671	9	US-10-982-512-2
14	2644.2	84.7	2649	6	US-10-054-776-1
15	2508.4	80.4	2830	3	US-09-956-157-2177
16	2500.4	80.1	4523	3	US-09-976-674-8
17	2500.4	80.1	4523	9	US-10-982-512-8
18	2208	70.8	2510	7	US-10-275-505-16
19	2208	70.8	2510	10	US-11-140-224-16
20	2056.4	65.9	4309	3	US-09-976-674-14
21	2056.4	65.9	4309	9	US-10-982-512-14
22	1347	43.2	1669	8	US-10-825-632-6
23	1036	33.2	1197	8	US-10-825-632-4

24	926.4	29.7	1356	3	US-09-976-674-10	Sequence 10, Appl
25	926.4	29.7	1356	7	US-10-982-512-10	Sequence 10, Appl
26	883.8	28.3	3287	7	US-10-415-122-3	Sequence 3, Appl
27	862.6	27.6	2660	7	US-10-072-012-223	Sequence 223, App
28	862.6	27.6	2660	7	US-10-072-012-225	Sequence 24, Appl
29	861	27.6	2617	3	US-09-976-674-4	Sequence 4, Appl
30	861	27.6	2617	9	US-10-982-512-4	Sequence 4, Appl
31	861	27.6	3716	9	US-10-433-757-30	Sequence 30, Appl
32	861	27.6	4219	3	US-09-976-674-28	Sequence 28, Appl
33	861	27.6	4219	9	US-10-982-512-28	Sequence 24, Appl
34	861	27.6	4302	3	US-09-976-674-24	Sequence 24, Appl
35	861	27.6	4302	9	US-10-982-512-24	Sequence 1, Appl
36	857.8	27.5	3000	7	US-10-415-122-1	Sequence 8, Appl
37	854	27.4	2495	7	US-10-415-122-8	Sequence 18, Appl
38	801.4	25.7	832	3	US-09-976-674-18	Sequence 18, Appl
39	801.4	25.7	832	9	US-10-982-512-18	Sequence 36, Appl
40	798.6	25.6	4180	3	US-09-976-674-36	Sequence 36, Appl
41	798.6	25.6	4180	9	US-10-982-512-36	Sequence 34, Appl
42	798.6	25.6	4263	3	US-09-976-674-34	Sequence 34, Appl
43	798.6	25.6	4263	9	US-10-982-512-34	Sequence 8, Appl
44	791.2	25.4	1083	8	US-10-825-632-8	Sequence 8, Appl
45	787.6	25.2	925	6	US-10-264-237-710	Sequence 710, Appl

ALIGNMENTS

RESULT 1
US-10-415-122-5
; Sequence 5, Application US/10415122
; Publication No. US20040053369A1
; GENERAL INFORMATION:
; APPLICANT: THE UNIVERSITY OF SYDNEY
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES
; FILE REFERENCE: P15217
; CURRENT APPLICATION NUMBER: US/10/415,122
; CURRENT FILING DATE: 2003-08-07
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 5
; LENGTH: 3120
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-415-122-5

Query Match 100.0%; Score 3120; DB 7; Length 3120;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	AAGTGTAAAGCTCCGAGGCCAAGCGCGCTGCTACTGCGCGCGCTGCTTCTTAGTGCGG	60
Db	1	AAGTGTAAAGCTCCGAGGCCAAGCGCGCTGCTACTGCGCGCGCTGCTTCTTAGTGCGG	60
Qy	61	CGTTGCGCGCGCTGGTTGTTCACCGCGCGCGCGCGAGAGCCACTCAACAGGACCG	120
Db	61	CGTTGCGCGCGCTGGTTGTTCACCGCGCGCGCGCGAGAGCCACTCAACAGGACCG	120
Qy	121	GAGTGGAGGCGCGGACGATGAAGCGCGCGCGCGCTCCATAGCGACGTCGGGACGG	180
Db	121	GAGTGGAGGCGCGGACGATGAAGCGCGCGCGCGCTCCATAGCGACGTCGGGACGG	180
Qy	181	TCCGCGCGCGCGCGGGAAGGAAATGCAACATGGCAGCAGCAATGGAACAGACAG	240
Db	181	TCCGCGCGCGCGCGGGAAGGAAATGCAACATGGCAGCAGCAATGGAACAGACAG	240
Qy	241	CTGGGTGTGAGATATTGAACTCGGACTGTGAGGAGAAATTTGAATTCACAGATCGG	300
Db	241	CTGGGTGTGAGATATTGAACTCGGACTGTGAGGAGAAATTTGAATTCACAGATCGG	300
Qy	301	CCTAAATGGAGCCCTTTTATTTGTTGAGCGGTATTCCTGAGTCAGCTTAAAGCTGCTT	360
Db	301	CCTAAATGGAGCCCTTTTATTTGTTGAGCGGTATTCCTGAGTCAGCTTAAAGCTGCTT	360

QY 361 GCCGATACCGAGAAATATCATGGCTACATGATGGCTAAGGCCACCATGATTTTCATGTTT 420
Db 361 GCCGATACCGAGAAATATCATGGCTACATGATGGCTAAGGCCACCATGATTTTCATGTTT 420
QY 421 GTGAAGAGGAATGATCCAGATGGACCTCAATTCAGACAGAACTCTATTACCTTGCCCATGTCT 480
Db 421 GTGAAGAGGAATGATCCAGATGGACCTCAATTCAGACAGAACTCTATTACCTTGCCCATGTCT 480
QY 481 GGTGAGAACAGAGAAATACACTGTTTTATTCTCGAAATTCACAACTATCAATAGAGCA 540
Db 481 GGTGAGAACAGAGAAATACACTGTTTTATTCTCGAAATTCACAACTATCAATAGAGCA 540
QY 541 GCAGCTTAAATGCTCTCTTGGAGCCTCTTTTGGATCTTTTTCAGCAACACTGCACTAT 600
Db 541 GCAGCTTAAATGCTCTCTTGGAGCCTCTTTTGGATCTTTTTCAGCAACACTGCACTAT 600
QY 601 GGAATGTAATCTCGAGAGAGAACTATTAAAGAGAGAAAGCAATGGACAGCTCGGA 660
Db 601 GGAATGTAATCTCGAGAGAGAACTATTAAAGAGAGAAAGCAATGGACAGCTCGGA 660
QY 661 ATTGCTTCTTAGGATTATCACCAAGGAAGTGGACATTTCTGTTTCAGCCGGTAGTGA 720
Db 661 ATTGCTTCTTAGGATTATCACCAAGGAAGTGGACATTTCTGTTTCAGCCGGTAGTGA 720
QY 721 ATTTATCATCGTAAAGATGGAGGCCCAAGGAATTTACGCAACAGCTTTAAAGGCCAAT 780
Db 721 ATTTATCATCGTAAAGATGGAGGCCCAAGGAATTTACGCAACAGCTTTAAAGGCCAAT 780
QY 781 CTAGTGGAACTAGTGTGCCCAACATACGGATGGATCCAAATATATGCCCCGCTGATCCA 840
Db 781 CTAGTGGAACTAGTGTGCCCAACATACGGATGGATCCAAATATATGCCCCGCTGATCCA 840
QY 841 GACTGGATTTGCTTTTATACATAGCAACGATATTTGGATATCTAACATCGTAAACAGAGAA 900
Db 841 GACTGGATTTGCTTTTATACATAGCAACGATATTTGGATATCTAACATCGTAAACAGAGAA 900
QY 901 GAAAGAGACTCACTTATGTGCACAATGAGCTAGGCCAACTGGAAGAAGATGCCAGATCA 960
Db 901 GAAAGAGACTCACTTATGTGCACAATGAGCTAGGCCAACTGGAAGAAGATGCCAGATCA 960
QY 961 GCTGGAGTCGCTTACCTTTGTTCTCCAAAGAGAAATTTGATAGATATCTGGCTATTGGTGG 1020
Db 961 GCTGGAGTCGCTTACCTTTGTTCTCCAAAGAGAAATTTGATAGATATCTGGCTATTGGTGG 1020
QY 1021 TGTCCAAAGCTGAAACAACTCCAGTGGTGAATTTCTTAGAATTTCTATATGAAGAA 1080
Db 1021 TGTCCAAAGCTGAAACAACTCCAGTGGTGAATTTCTTAGAATTTCTATATGAAGAA 1080
QY 1081 AATGATGATCTGAGGTGGAAATTTATCATGTTACATCCCTATGTTGGAACAGAGG 1140
Db 1081 AATGATGATCTGAGGTGGAAATTTATCATGTTACATCCCTATGTTGGAACAGAGG 1140
QY 1141 GCAGATTCATTCGGTTATCCTTAAACAGGTTACAGCAATCCTTAAAGTCACTTTTAAGATG 1200
Db 1141 GCAGATTCATTCGGTTATCCTTAAACAGGTTACAGCAATCCTTAAAGTCACTTTTAAGATG 1200
QY 1201 TCAGAAATATGATTTGATGCTGAAGGAAGGATCATAGATGCTATAGATAAGGAATTAAT 1260
Db 1201 TCAGAAATATGATTTGATGCTGAAGGAAGGATCATAGATGCTATAGATAAGGAATTAAT 1260
QY 1261 CAACCTTTTGAGATCTATTGGAAGGATTTGAATATATTCAGAGCTGGATGCACTCT 1320
Db 1261 CAACCTTTTGAGATCTATTGGAAGGATTTGAATATATATTCAGAGCTGGATGCACTCT 1320
QY 1321 GAGGAAATATGCTTGGTCCATCCTACTAGATGGCTCCAGACTCGCTTACAGATAGTG 1380
Db 1321 GAGGAAATATGCTTGGTCCATCCTACTAGATGGCTCCAGACTCGCTTACAGATAGTG 1380
QY 1381 TTGATCTCACCTGGAATTTATTTATCCAGTAGAAGATGATGTTATGGAAGGAGAGACTC 1440
Db 1381 TTGATCTCACCTGGAATTTATTTATCCAGTAGAAGATGATGTTATGGAAGGAGAGACTC 1440
QY 1441 ATTGAGTCAGTGCCTGATTTCTGTGACGCCACTAATTTATCTATGAAGAAACACAGACATC 1500

Db 1441 ATTGAGTCAGTGCCTGATTTCTGTGACGCCACTAATTTATCTATGAAGAAACACAGACATC 1500
QY 1501 TGGATAAATATCCATGACATCTTTTCATGTTTTCCTCCAAAGTACAGAGAGAAATTGAG 1560
Db 1501 TGGATAAATATCCATGACATCTTTTCATGTTTTCCTCCAAAGTACAGAGAGAAATTGAG 1560
QY 1561 TTTATTTTGGCTCTGAAATGCAAAAGTTTCCGTCTATTATACAAAATTACATCTATT 1620
Db 1561 TTTATTTTGGCTCTGAAATGCAAAAGTTTCCGTCTATTATACAAAATTACATCTATT 1620
QY 1621 TTTAAAGGAAACAAATATAAACGATCCAGTGGTGGGCTGCTCTCAAAGTGAATTTCAAG 1680
Db 1621 TTTAAAGGAAACAAATATAAACGATCCAGTGGTGGGCTGCTCTCAAAGTGAATTTCAAG 1680
QY 1681 TGTCTCTATCAAAGAGGAGATAGCAATTTACCAGTGGTGAATGGGAAGTTCTTGGCCGGCAT 1740
Db 1681 TGTCTCTATCAAAGAGGAGATAGCAATTTACCAGTGGTGAATGGGAAGTTCTTGGCCGGCAT 1740
QY 1741 GGATCTAATATCCAAGTTGATGAAGTCAAGAGGCTGGTATATTTTGAAGGCAACAAAGAC 1800
Db 1741 GGATCTAATATCCAAGTTGATGAAGTCAAGAGGCTGGTATATTTTGAAGGCAACAAAGAC 1800
QY 1801 TCCCTTTTAGAGCATCACTGTACGTAGTCACTTCTGTGCATCAGTCAGCATCTGCTTTTATA 1860
Db 1801 TCCCTTTTAGAGCATCACTGTACGTAGTCACTTCTGTGCATCAGTCAGCATCTGCTTTTATA 1860
QY 1861 CTGACTGACCGTGGCTACTCACATTTCTGTCATCAGTCAGCATCTGCTTTTATA 1920
Db 1861 CTGACTGACCGTGGCTACTCACATTTCTGTCATCAGTCAGCATCTGCTTTTATA 1920
QY 1921 AGTAAGTATAGTAACCAAGAAATCCACACTGTTGTCCTTTTACAAAGCTATCAAGTCT 1980
Db 1921 AGTAAGTATAGTAACCAAGAAATCCACACTGTTGTCCTTTTACAAAGCTATCAAGTCT 1980
QY 1981 GAAAGTACCCAACTTTCGAAAACAAAGAAATTTTGGCCACCATTTTGGATTCAGCAGGT 2040
Db 1981 GAAAGTACCCAACTTTCGAAAACAAAGAAATTTTGGCCACCATTTTGGATTCAGCAGGT 2040
QY 2041 CCTTCTCTGACTATACCTCTCCAGAAATTTTCTTTTGAAGTACTACTGGAATTTACA 2100
Db 2041 CCTTCTCTGACTATACCTCTCCAGAAATTTTCTTTTGAAGTACTACTGGAATTTACA 2100
QY 2101 TTGTATGGGATGCTCTACAAGCCTCATGATCTACAGCCTCGAAAGAAATATCTACTGTG 2160
Db 2101 TTGTATGGGATGCTCTACAAGCCTCATGATCTACAGCCTCGAAAGAAATATCTACTGTG 2160
QY 2161 CTGTTTCATATATGGTGGTCTCTCAGTGCAGTGGTGAATTAATCGGTTTAAAGGAGTCAAG 2220
Db 2161 CTGTTTCATATATGGTGGTCTCTCAGTGCAGTGGTGAATTAATCGGTTTAAAGGAGTCAAG 2220
QY 2221 TATTTCCGCTTGAATACCTTAGCCTCTCTAGGTTATGTGGTGTAGTGTAGTGAACACAGG 2280
Db 2221 TATTTCCGCTTGAATACCTTAGCCTCTCTAGGTTATGTGGTGTAGTGTAGTGAACACAGG 2280
QY 2281 GGATCTGTCACCGAGGGCTTAAATTTGAAGGCCCTTTTAAATATAAAATGGGTCAAATA 2340
Db 2281 GGATCTGTCACCGAGGGCTTAAATTTGAAGGCCCTTTTAAATATAAAATGGGTCAAATA 2340
QY 2341 GAAATTCAGATCAGTGGAGGACTCCAATATCTAGCTTCTCGATGATGATTTCAATTGAC 2400
Db 2341 GAAATTCAGATCAGTGGAGGACTCCAATATCTAGCTTCTCGATGATGATTTCAATTGAC 2400
QY 2401 TTAGATCTGTGGGATTCACCGGCTGGTCTTATGGAAGATCCTCTCCCTGGGCAATTA 2460
Db 2401 TTAGATCTGTGTGGGATTCACCGGCTGGTCTTATGGAAGATCCTCTCCCTGGGCAATTA 2460
QY 2461 ATCCAGAGTTCAGATATCTTTCAGGGTTCTTATGCGGGGCCCCAGTCACTCTGTGGATC 2520
Db 2461 ATCCAGAGTTCAGATATCTTTCAGGGTTCTTATGCGGGGCCCCAGTCACTCTGTGGATC 2520
QY 2521 TTCTATGATACAGGATACACGGAACTTTATATGGGTCACCCCTGACCGAGATGAACAGGGC 2580

2521	TTCTATGATACAGGATACACGGAAAGTTATATGGGTCACTCCGTGACCAAGATGAACAGGGC	2588
Db		
2581	TATTACTTTAGGATCTGTGGCCATGCAAGACGAAAAGTTCCCTCTGTGAACCAAAATCGTTTA	2640
Qy		
2581	TATTACTTTAGGATCTGTGGCCATGCAAGACGAAAAGTTCCCTCTGTGAACCAAAATCGTTTA	2640
Db		
2641	CTGTCTTTACATGGTTTCTCGGATGAGAATGTCCATTTTGGACATACACAGTATATTACTG	2700
Qy		
2641	CTGTCTTTACATGGTTTCTCGGATGAGAATGTCCATTTTGGACATACACAGTATATTACTG	2700
Db		
2701	AGTTTTTTAGTGAAGGCTGGAAAGCCCATATGATTTACAGATCTATCTCTCAGAGAGACAC	2760
Qy		
2701	AGTTTTTTAGTGAAGGCTGGAAAGCCCATATGATTTACAGATCTATCTCTCAGAGAGACAC	2760
Db		
2761	AGCATAGAAGTTCTCGAATCGGAGAACATATGAAGACTGCATCTTTTGGCACTACCTTCAA	2820
Qy		
2761	AGCATAGAAGTTCTCGAATCGGAGAACATATGAAGACTGCATCTTTTGGCACTACCTTCAA	2820
Db		
2821	GAAGAACCTTGGATCAACGTAATGGCTCTTAAAGTGATATAATTTTGGACCTGTGTAGAAC	2880
Qy		
2821	GAAGAACCTTGGATCAACGTAATGGCTCTTAAAGTGATATAATTTTGGACCTGTGTAGAAC	2880
Db		
2881	TCTCTGGTATACATGCGGCTATTAAACCAATGAGGAGGTTTAATCAACAGAAAACACAGA	2940
Qy		
2881	TCTCTGGTATACATGCGGCTATTAAACCAATGAGGAGGTTTAATCAACAGAAAACACAGA	2940
Db		
2941	ATTGATCATCATTTTGTATACCTGACATGCAATCTACTCCTGAAATAAATGTTGGT	3000
Qy		
2941	ATTGATCATCATTTTGTATACCTGACATGCAATCTACTCCTGAAATAAATGTTGGT	3000
Db		
3001	CCATGACGGGGTCTACGGTTTGTGGTAGTAATCTTAATACCTTAAACCCACATGCTCAAAA	3060
Qy		
3001	CCATGACGGGGTCTACGGTTTGTGGTAGTAATCTTAATACCTTAAACCCACATGCTCAAAA	3060
Db		
3061	TCAATATGATACATATTCTCTGAGAGACCCAGCAATACCATAGAATTTACTTAAAAAATAAAA	3120
Qy		
3061	TCAATATGATACATATTCTCTGAGAGACCCAGCAATACCATAGAATTTACTTAAAAAATAAAA	3120
Db		

RESULT 2

```

US-10-825-632-2
; Sequence 2, Application US/10825632
; Publication No. US20040191826A1
; GENERAL INFORMATION:
; APPLICANT: ABBOTT, Catherine Anne
; APPLICANT: CORRELL, Mark Douglas
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES
; FILE REFERENCE: FCSB-100-Div. 1
; CURRENT APPLICATION NUMBER: US/10/825,632
; CURRENT FILING DATE: 2004-04-15
; PRIOR APPLICATION NUMBER: US 10/070,464
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: PCT/AU00/01085
; PRIOR FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: AU PQ5709
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: AU PQ2762
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 3120
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-825-632-2

```

[illegible]

1141 GCAGATTCAATCCGTTATCTTAACAGGTAAGCAATCCCTAAGTCACTTTTAAGATG 1200
1141 GCAGATTCAATCCGTTATCTTAACAGGTAAGCAATCCCTAAGTCACTTTTAAGATG 1200
1201 TCAGAAATAATGATGATGCTGAAGAGGATCATAGATGTCATAGATAAGGAATTAAT 1260
1201 TCAGAAATAATGATGATGCTGAAGAGGATCATAGATGTCATAGATAAGGAATTAAT 1260
1261 CAACCTTTTGAGATCTTAATTTGAAGGAGTTGAATATATTTGCCAGAGCTGGATGCTCT 1320
1261 CAACCTTTTGAGATCTTAATTTGAAGGAGTTGAATATATATTTGCCAGAGCTGGATGCTCT 1320
1321 GAGGAAAATAATGCTTGCTGCTCACTACTAGATCGCTCCAGAGCTGGATGCTGGATG 1380
1321 GAGGAAAATAATGCTTGCTGCTCACTACTAGATCGCTCCAGAGCTGGATGCTGGATG 1380
1381 TTGATCTCACTCACTGAATATTTATCCAGTAGAAGATGATGTTATGGAAGGAGAGAGCTC 1440
1381 TTGATCTCACTCACTGAATATTTATCCAGTAGAAGATGATGTTATGGAAGGAGAGAGCTC 1440
1441 ATTGAGTCAGTGCCTGATCTCTGACGCCACTAATATCTATGAAGAAACACAGACATC 1500
1441 ATTGAGTCAGTGCCTGATCTCTGACGCCACTAATATCTATGAAGAAACACAGACATC 1500
1501 TGGATAAATAATCCATGACATCTTTCACTGTTTCCCAAGTCAAGAGGAGAAATTCAG 1560
1501 TGGATAAATAATCCATGACATCTTTCACTGTTTCCCAAGTCAAGAGGAGAAATTCAG 1560
1561 TTTATTTTGGCTCTGAAATGCAAAACAGGTTTCCGTCATTTATACAAAATTAACATCTATT 1620
1561 TTTATTTTGGCTCTGAAATGCAAAACAGGTTTCCGTCATTTATACAAAATTAACATCTATT 1620
1621 TTTAAAGGAAGCAATATAAACGATCCAGTGGTGGCTGCTCTCCAAAGTATTCAAG 1680
1621 TTTAAAGGAAGCAATATAAACGATCCAGTGGTGGCTGCTCTCCAAAGTATTCAAG 1680
1681 TGCTCTATCAAGAGGAGATAGCAATTAACAGTGGTGAATGGGAAGTTCTTGCGCGGCAT 1740
1681 TGCTCTATCAAGAGGAGATAGCAATTAACAGTGGTGAATGGGAAGTTCTTGCGCGGCAT 1740
1741 GGATCTAATATCCAAAGTTGATGAAGTCAGAAAGGCTGGTATATTTGAAGGCCACCAAGAC 1800
1741 GGATCTAATATCCAAAGTTGATGAAGTCAGAAAGGCTGGTATATTTGAAGGCCACCAAGAC 1800
1801 TCCCTTTTAGAGCATCACTCTGATAGTCACTAGTCACTAGTCACTGAGAGGTGACAGG 1860
1801 TCCCTTTTAGAGCATCACTCTGATAGTCACTAGTCACTAGTCACTGAGAGGTGACAGG 1860
1861 CTGACTGACCGTGGTACTACATCTTCTGCGATCAGTCACTGATCTGATCTTTTATA 1920
1861 CTGACTGACCGTGGTACTACATCTTCTGCGATCAGTCACTGATCTGATCTTTTATA 1920
1921 AGTAAGTATAGTAACAGAGAAATCCACATGCTGTGCTCTTTTCAAGCTATCAAGTCTCT 1980
1921 AGTAAGTATAGTAACAGAGAAATCCACATGCTGTGCTCTTTTCAAGCTATCAAGTCTCT 1980
1981 GAAGATGACCCCACTTGCAAAACAAAGGAATTTTGGCCACCAATTTTGGATTGAGAGT 2040
1981 GAAGATGACCCCACTTGCAAAACAAAGGAATTTTGGCCACCAATTTTGGATTGAGAGT 2040
2041 CCTCTCTGACTACTACTCTCCAGAAATTTCTCTTTTGAAGTACTACTGATTTACA 2100
2041 CCTCTCTGACTACTACTCTCCAGAAATTTCTCTTTTGAAGTACTACTGATTTACA 2100
2101 TTGATGAGATCTCTACAGCCTCATGATCTACAGCCTGGAAGAAATATCTACTGTG 2160
2101 TTGATGAGATCTCTACAGCCTCATGATCTACAGCCTGGAAGAAATATCTACTGTG 2160
2161 CTGTTTCATATAGTGGTCTCAGTGCAGTGGTGAATATCGTTTAAAGGAGTCAAG 2220
2161 CTGTTTCATATAGTGGTCTCAGTGCAGTGGTGAATATCGTTTAAAGGAGTCAAG 2220
2221 TATTTCCGCTTGAATACCTAGCCTCTCTAGGTTATGCTGGTTAGTAGACACAGG 2280

2221 TATTTCCGCTTGAATACCTAGCCTCTCTAGGTTATGCTGGTTAGTAGACACAGG 2280
2281 GGAATCTGTCAACGAGGCTTTAAATTTGAAGGCGCTTTAAATATATAAATGGGTCAATA 2340
2281 GGAATCTGTCAACGAGGCTTTAAATTTGAAGGCGCTTTAAATATATAAATGGGTCAATA 2340
2341 GAAATGACGATCAGGTGGAGGATCCCAATATCTAGCTTCTCGATATGATTTTCATGAC 2400
2341 GAAATGACGATCAGGTGGAGGATCCCAATATCTAGCTTCTCGATATGATTTTCATGAC 2400
2401 TTAGATCGTGTGGGCATCCACGGCTGGTCTATGAGGATACCTCTCCCTGATGAGGATTA 2460
2401 TTAGATCGTGTGGGCATCCACGGCTGGTCTATGAGGATACCTCTCCCTGATGAGGATTA 2460
2461 ATGCAAGAGTCAGATATCTTCAGGGTTGCTATTGCTGGGGCCCGAGTCACTCTGTGGATC 2520
2461 ATGCAAGAGTCAGATATCTTCAGGGTTGCTATTGCTGGGGCCCGAGTCACTCTGTGGATC 2520
2521 TTTCTATGATACAGGATACACGGAACGTTATATGGGTCACTCTGACAGAAATGAACAGGGC 2580
2521 TTTCTATGATACAGGATACACGGAACGTTATATGGGTCACTCTGACAGAAATGAACAGGGC 2580
2581 TATTACTTAGGATCTGTGGCATGCAAGCAGAAAAGTTCCCTCTGAAACCAAAATCGTTTA 2640
2581 TATTACTTAGGATCTGTGGCATGCAAGCAGAAAAGTTCCCTCTGAAACCAAAATCGTTTA 2640
2641 CTGCTCTTACATGCTTCCGATGAGATGTCATTTTGCACATACCACTATATTACTG 2700
2641 CTGCTCTTACATGCTTCCGATGAGATGTCATTTTGCACATACCACTATATTACTG 2700
2701 AGTTTTTATGAGGCTGGAAAGCCATATGATTTACAGATCTATCTCAGGAGACAC 2760
2701 AGTTTTTATGAGGCTGGAAAGCCATATGATTTACAGATCTATCTCAGGAGACAC 2760
2761 AGCATAAGAGTTTCTGAAATCCGGAGAACATPATGAACTGCACTTTTGGCACTACCTTCAA 2820
2761 AGCATAAGAGTTTCTGAAATCCGGAGAACATPATGAACTGCACTTTTGGCACTACCTTCAA 2820
2821 GAAACCTTTGGATCAGCTATTGCTGCTCTAAAGTGATATAAATTTGACCTGTGAGAAC 2880
2821 GAAACCTTTGGATCAGCTATTGCTGCTCTAAAGTGATATAAATTTGACCTGTGAGAAC 2880
2881 TCTCTGGTATACACTGGCTATTAAACCAAAATGAGGAGGTTTAAATCAACAGAAACACAGA 2940
2881 TCTCTGGTATACACTGGCTATTAAACCAAAATGAGGAGGTTTAAATCAACAGAAACACAGA 2940
2941 ATTGATCATCATATTTGATACCTGCGCATGTAACATCTACTCTGAAAAATAAATGTGGTG 3000
2941 ATTGATCATCATATTTGATACCTGCGCATGTAACATCTACTCTGAAAAATAAATGTGGTG 3000
3001 CCATGCGGGTCTACGGTTTGGTGTAGTATCTAATACCTTAACCCCATGCTCAAAA 3060
3001 CCATGCGGGTCTACGGTTTGGTGTAGTATCTAATACCTTAACCCCATGCTCAAAA 3060
3061 TCAATGATACATATTTCTGAGAGCCCGCAATACCTAAGATTTACTAAAAAATAA 3120
3061 TCAATGATACATATTTCTGAGAGCCCGCAATACCTAAGATTTACTAAAAAATAA 3120

RESULT 3

US-09-976-674-12
; Sequence 12, Application US/09976674
; Patent No. US20020115843A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12

;; PRIOR APPLICATION NUMBER: US 60/240,117
;; PRIOR FILING DATE: 2000-10-12
;; NUMBER OF SEQ ID NOS: 61
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 12
;; LENGTH: 4829
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-09-976-674-12

Query Match		99.6%	Score 3106.4;	DB 3;	Length 4829;
Best Local Similarity		99.9%	Pred. No. 0;		
Matches 3119;		Conservative	1;	Indels	2; Gaps 1;
Qy	1	AAGTGCTAAGAGCTCCGAGGCCCAAGGCCGCTGCTACTGCGCGCGCTCTTCTTAGTGCCG	60		
Dd	1	AAGTGCTAAGAGCTCCGAGGCCCAAGGCCGCTGCTACTGCGCGCGCTCTTCTTAGTGCCG	60		
Qy	61	CGTTTCGCGCGCTGGGTTGTCAACCGCGCGCGCGCGCGAGGAGCCACTGCAACCCAGGACCG	120		
Dd	61	CGTTTCGCGCGCTGGGTTGTCAACCGCGCGCGCGCGCGAGGAGCCACTGCAACCCAGGACCG	120		
Qy	121	GAGTGGAGCGCGCGAGCATGAAGCGCGAGCGCGCGCTCCATAGCGCACGTCGCGACCG	180		
Dd	121	GAGTGGAGCGCGCGAGCATGAAGCGCGAGCGCGCGCTCCATAGCGCACGTCGCGACCG	180		
Qy	181	TCCGGCGCGCGCGGGGAGGAAATGCAATGCGCAGCAGCAATGGAACAGAACAG	240		
Dd	181	TCCGGCGCGCGCGGGGAGGAAATGCAATGCGCAGCAGCAATGGAACAGAACAG	240		
Qy	241	CTGGGTGTTGAGATATTTGAAACTGCGGACTGTGAGGAGAAATTTGAATCACAGGATCGG	300		
Dd	241	CTGGGTGTTGAGATATTTGAAACTGCGGACTGTGAGGAGAAATTTGAATCACAGGATCGG	300		
Qy	301	CCTAAATTGGAGCCTTTTATGTTGAGCGGTATTCCTGAGTCAGCTTAAAGTGTCTT	360		
Dd	301	CCTAAATTGGAGCCTTTTATGTTGAGCGGTATTCCTGAGTCAGCTTAAAGTGTCTT	360		
Qy	361	GCCGATACAGAAATATCATGCTACATGCTAGCTAGCGTAAAGCACCACATGATTTTCA	420		
Dd	361	GCCGATACAGAAATATCATGCTAGCTAGCGTAAAGCACCACATGATTTTCA	420		
Qy	421	GTGAAGAGGAATGATCCAGATGGACCTCATTTCAGACAGAACTTATTAACCTGCCATGCT	480		
Dd	421	GTGAAGAGGAATGATCCAGATGGACCTCATTTCAGACAGAACTTATTAACCTGCCATGCT	480		
Qy	481	GGTGAGAACAGAGAAATACACTGTTTATTCGAAATTCGAAATTCGAAATTCGAAATTCG	540		
Dd	481	GGTGAGAACAGAGAAATACACTGTTTATTCGAAATTCGAAATTCGAAATTCGAAATTCG	540		
Qy	541	GCAGTCTTAATGCTCTCTTGAAGCCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT	600		
Dd	541	GCAGTCTTAATGCTCTCTTGAAGCCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT	600		
Qy	601	GGAATGATATTCGAGAGAGAACTATTAAGAGAAAGAAACGCAATGGACAGTCGGA	660		
Dd	601	GGAATGATATTCGAGAGAGAACTATTAAGAGAAAGAAACGCAATGGACAGTCGGA	660		
Qy	661	ATTGCTTCTTACGATTATCACCAAGGAAGTGGAACTTTCTGTTTCAAGCCGGTAGTGA	720		
Dd	661	ATTGCTTCTTACGATTATCACCAAGGAAGTGGAACTTTCTGTTTCAAGCCGGTAGTGA	720		
Qy	721	ATTATACGTTAAAGATGGAGGCCACAGGATTTACGCAACAACTTTTAAGGCCCAAT	780		
Dd	721	ATTATACGTTAAAGATGGAGGCCACAGGATTTACGCAACAACTTTTAAGGCCCAAT	780		
Qy	781	CTAGTGGAAATAGTTGTCCCAACATACGATGGATCCAAATTTATGCCCGCTGATCCA	840		
Dd	781	CTAGTGGAAATAGTTGTCCCAACATACGATGGATCCAAATTTATGCCCGCTGATCCA	840		
Qy	841	GACTGGATTGCTTTTATACATAGCAACGATATTTGGATATCTAACATCGTAAACAGAGAA	900		
Dd	841	GACTGGATTGCTTTTATACATAGCAACGATATTTGGATATCTAACATCGTAAACAGAGAA	900		

Qy	901	GAAGGAGAGCTCACTTATGTGCAATGAGCTAGCCAAACATCGAAGAGATGCCAGATCA	960		
Dd	901	GAAGGAGAGCTCACTTATGTGCAATGAGCTAGCCAAACATCGAAGAGATGCCAGATCA	960		
Qy	961	GCTGGAGTCGCTACCTTTGTTCTCCAAAGAAATTTGATAGATATTTCTGGCTATTGGTGG	1020		
Dd	961	GCTGGAGTCGCTACCTTTGTTCTCCAAAGAAATTTGATAGATATTTCTGGCTATTGGTGG	1020		
Qy	1021	TGTCAAAAGCTGAAACAACTCCCAAGTGGTGGTAAATTTCTTAGAATCTCTATGAAGAA	1080		
Dd	1021	TGTCAAAAGCTGAAACAACTCCCAAGTGGTGGTAAATTTCTTAGAATCTCTATGAAGAA	1080		
Qy	1081	AATGATGATCTGAGGTGGAAATTTATCATGTTTACATCCCTATGTTGGAACAAGGAGG	1140		
Dd	1081	AATGATGATCTGAGGTGGAAATTTATCATGTTTACATCCCTATGTTGGAACAAGGAGG	1140		
Qy	1141	GCAGATTCATTTCCGTTATCTTAAACAGGTACAGCAATCTTAAAGTCACTTTTAAAGATG	1200		
Dd	1141	GCAGATTCATTTCCGTTATCTTAAACAGGTACAGCAATCTTAAAGTCACTTTTAAAGATG	1200		
Qy	1201	TCAGAAATAATGATGATGCTGAAAGGAAGATCATAGATGCTATAGATAGGAACCTAAT	1260		
Dd	1201	TCAGAAATAATGATGATGCTGAAAGGAAGATCATAGATGCTATAGATAGGAACCTAAT	1260		
Qy	1261	CAACCTTTTGAGATTTCTATTTGAAGGAGTTGAATATATTTGCCAGAGCTGGATGCTCT	1320		
Dd	1261	CAACCTTTTGAGATTTCTATTTGAAGGAGTTGAATATATTTGCCAGAGCTGGATGCTCT	1320		
Qy	1321	GAGGGAATAATGCTTGGTCCATCTACTAGATCGCTCCAGAGCTCGCTTACAGATAGTG	1380		
Dd	1321	GAGGGAATAATGCTTGGTCCATCTACTAGATCGCTCCAGAGCTCGCTTACAGATAGTG	1380		
Qy	1381	TTGATCTCACCTGAAATTTATTTCCAGTAGAAGATCATGTTATGGAAGAGCAGAGACTC	1440		
Dd	1381	TTGATCTCACCTGAAATTTATTTCCAGTAGAAGATCATGTTATGGAAGAGCAGAGACTC	1440		
Qy	1441	ATTGAGTCAGTCCTGATTTCTGACGCCACTAATTTATCTATGAAGAAACAAACAGACTC	1500		
Dd	1441	ATTGAGTCAGTCCTGATTTCTGACGCCACTAATTTATCTATGAAGAAACAAACAGACTC	1500		
Qy	1501	TGGAATAATATCCATGACATCTTTTCATGTTTCCCAAAGTCCAGAAAGGAAATTTGAG	1560		
Dd	1501	TGGAATAATATCCATGACATCTTTTCATGTTTCCCAAAGTCCAGAAAGGAAATTTGAG	1560		
Qy	1561	TTTATTTTGGCTCTGAAATGCAAAACAGGTTTCCGTCATTATACAAAATTTACATCTAT	1620		
Dd	1561	TTTATTTTGGCTCTGAAATGCAAAACAGGTTTCCGTCATTATACAAAATTTACATCTAT	1620		
Qy	1621	TTAAAGGAAAGCAAAATATAAACGATCCAGTGGTGGCTGCTCCAAAGTATTTCAAG	1680		
Dd	1621	TTAAAGGAAAGCAAAATATAAACGATCCAGTGGTGGCTGCTCCAAAGTATTTCAAG	1680		
Qy	1681	TGTCCTCTCAAGAGAGAGATAGCAATTAACAGTGGTGGCTGCTCCAAAGTATTTCAAG	1740		
Dd	1681	TGTCCTCTCAAGAGAGAGATAGCAATTAACAGTGGTGGCTGCTCCAAAGTATTTCAAG	1740		
Qy	1741	GGATCTAAATATCCAAAGTGGATGAGTCAAGAGGCTGGTATATTTTGAAGGCAACAAAG	1800		
Dd	1741	GGATCTAAATATCCAAAGTGGATGAGTCAAGAGGCTGGTATATTTTGAAGGCAACAAAG	1800		
Qy	1801	TCCCTCTTATAGATCACTGATGATGATGATGATGATGATGATGATGATGATGATGATG	1860		
Dd	1801	TCCCTCTTATAGATCACTGATGATGATGATGATGATGATGATGATGATGATGATGATG	1860		
Qy	1861	CTGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1920		
Dd	1861	CTGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1920		
Qy	1921	AGTAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1980		
Dd	1921	AGTAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1980		

QY 1981 GAAGATGACCAACTTGCAGAAACAAAGGAATTTTGGGCCACCAATTTTGGATTGACAGGT 2040
DB 1981 GAAGATGACCAACTTGCAGAAACAAAGGAATTTTGGGCCACCAATTTTGGATTGACAGGT 2040
QY 2041 CCTCTTCTGACTACTACTCTCTCAGAAATTTTCTCTTTTGAAGTACTACTGATTTTACA 2100
DB 2041 CCTCTTCTGACTACTACTCTCTCAGAAATTTTCTCTTTTGAAGTACTACTGATTTTACA 2100
QY 2101 TTGTATGGGATGCTCTACAGCCTCATGATCTACAGCCTGGGAAGAAATATCTTACTGTG 2160
DB 2101 TTGTATGGGATGCTCTACAGCCTCATGATCTACAGCCTGGGAAGAAATATCTTACTGTG 2160
QY 2161 CTGTTTCAATATGTTGG--TCCTCAGGTGAGTTGGTGAATATCGGTTTAAAGAGTCA 2218
DB 2161 CTGTTTCAATATGTTGGTCTCTCAGGTGAGTTGGTGAATATCGGTTTAAAGAGTCA 2220
QY 2219 AGTATTTCCGTTGAAACCTTACCTTCTAGCTCTCTAGGTTATGTTGGTGGAGTATAGACA 2278
DB 2221 AGTATTTCCGTTGAAACCTTACCTTCTAGCTCTCTAGGTTATGTTGGTGGAGTATAGACA 2280
QY 2279 GGGGATCTGTCAACGAGGGCTTAAATTTGAAGGCGCTTTAAATATATAAAATGGGTCAA 2338
DB 2281 GGGGATCTGTCAACGAGGGCTTAAATTTGAAGGCGCTTTAAATATATAAAATGGGTCAA 2340
QY 2339 TAGAAATTTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTTCA 2398
DB 2341 TAGAAATTTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTTCA 2400
QY 2399 ACTTAGATCGTGTGGGATCCAGGCTGGTCTTATGAGGATACCTCTCCCTGATGGCAT 2458
DB 2401 ACTTAGATCGTGTGGGATCCAGGCTGGTCTTATGAGGATACCTCTCCCTGATGGCAT 2460
QY 2459 TAATGCCAGATCAGATATCTTCAGGTTGCTATCTGGGCGCCAGTCACTGTGGA 2518
DB 2461 TAATGCCAGATCAGATATCTTCAGGTTGCTATCTGGGCGCCAGTCACTGTGGA 2520
QY 2519 TCTTCTATGATACAGGATACCGGAACGTTTATATGTTGTTGCTGACCGAATGAACAGG 2578
DB 2521 TCTTCTATGATACAGGATACCGGAACGTTTATATGTTGTTGCTGACCGAATGAACAGG 2580
QY 2579 GCTATTTACTTAGATCTGTGGCCAAGCAAGCAAGAAAGTTCCCTCTGAAACCAATCGTT 2638
DB 2581 GCTATTTACTTAGATCTGTGGCCAAGCAAGCAAGAAAGTTCCCTCTGAAACCAATCGTT 2640
QY 2639 TACTGCTCTTACATGTTTCTGATGAGAAATGTCATTTTGACATACCAATGATATTAC 2698
DB 2641 TACTGCTCTTACATGTTTCTGATGAGAAATGTCATTTTGACATACCAATGATATTAC 2700
QY 2699 TGAGTTTCTTACATGTTTCTGATGAGAAATGTCATTTTGACATACCAATGATATTAC 2758
DB 2701 TGAGTTTCTTACATGTTTCTGATGAGAAATGTCATTTTGACATACCAATGATATTAC 2760
QY 2759 ACAGCATAGAGTTTCTGAAATCGGAGAACATTTATGAATGATCTATCTCTCAGGAGAGAC 2818
DB 2761 ACAGCATAGAGTTTCTGAAATCGGAGAACATTTATGAATGATCTATCTCTCAGGAGAGAC 2820
QY 2819 AAGAAACCTTGGATCAGTATTTGCTGCTTAAAGTGAATAATTTTGACCTGTGAGA 2878
DB 2821 AAGAAACCTTGGATCAGTATTTGCTGCTTAAAGTGAATAATTTTGACCTGTGAGA 2880
QY 2879 ACTCTGCTGATACACCTGGCTATTTAACCAATGAGGAGTTTATCAACAGAAACACA 2938
DB 2881 ACTCTGCTGATACACCTGGCTATTTAACCAATGAGGAGTTTATCAACAGAAACACA 2940
QY 2939 GAATGATCATCATTTTGTGATACCTGCCATGTAATCTACTCTGAAATTAATATGTTG 2998
DB 2941 GAATGATCATCATTTTGTGATACCTGCCATGTAATCTACTCTGAAATTAATATGTTG 3000
QY 2999 TGCCATGCGGGGTCTACGGTTTGTGTTAGTAACTTAATACTTAACCCCACTAGCTCAA 3058
DB 3001 TGCCATGCGGGGTCTACGGTTTGTGTTAGTAACTTAATACTTAACCCCACTAGCTCAA 3060
QY 3059 AATCAATGATACATATTTCTGAGAGACCCAGCAATACCATTAAGAAATTTACTAAAAA 3118

DB 3061 AATCAATGATACATATTTCTGAGAGACCCAGCAATACCATNAGAAATTTACTAAAAA 3120
QY 3119 AA 3120
DB 3121 AA 3122
RESULT 4
US-10-982-512-12
; Sequence 12, Application US/10982512
; Publication No. US20050059081A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Jean-Louis
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPP1V
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/10/982,512
; PRIOR FILING DATE: 2004-11-05
; PRIOR APPLICATION NUMBER: US/09/976,674
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 4829
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-982-512-12
Query Match 99.6%; Score 3106.4; DB 9; Length 4829;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3119; Conservative 0; Mismatches 1; Indels 2; Gaps 1;
QY 1 AAGTCTAAAGCTCCGAGGCCAAGCGCTGTCTACTGCGCGCTGCTTTTAGTGGCG 60
DB 1 AAGTCTAAAGCTCCGAGGCCAAGCGCTGTCTACTGCGCGCTGCTTTTAGTGGCG 60
QY 61 CGTTCCGCGCTGGGTTGTCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
DB 61 CGTTCCGCGCGCTGGGTTGTCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
QY 121 GAGTGGAGCGCGCGCGCGCATGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
DB 121 GAGTGGAGCGCGCGCGCGCATGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180
QY 181 TCCGGCG 240
DB 181 TCCGGCG 240
QY 241 CTGGGTGTTGAGATATTTGAACTGCGGACTGTGAGGAGAAATTTGAATCAGAGGATCG 300
DB 241 CTGGGTGTTGAGATATTTGAACTGCGGACTGTGAGGAGAAATTTGAATCAGAGGATCG 300
QY 301 CCTAAATTTGAGCGCTTTTATGTTGAGCGGTATTTCTGGAGTCAGCTTAAAGCTGCTT 360
DB 301 CCTAAATTTGAGCGCTTTTATGTTGAGCGGTATTTCTGGAGTCAGCTTAAAGCTGCTT 360
QY 361 GCCGATACAGAAATATCATGGCTCATGATGGCTAAGGCAACCATGATTTTCTGTTT 420
DB 361 GCCGATACAGAAATATCATGGCTCATGATGGCTAAGGCAACCATGATTTTCTGTTT 420
QY 421 GTGAAGAGAAATGATTCAGATGGAACCTTCTTACAGACAGAAATTTATCTTGCATGTCT 480
DB 421 GTGAAGAGAAATGATTCAGATGGAACCTTCTTACAGACAGAAATTTATCTTGCATGTCT 480
QY 481 GGTGAGAAACAGAGAAATACATCTGTTTATTTCTGAAATTTCCCAAACTATCAATAGACA 540
DB 481 GGTGAGAAACAGAGAAATACATCTGTTTATTTCTGAAATTTCCCAAACTATCAATAGACA 540

541 GCAGTCTTAATGCTCTCTTGGGAAGCCTCTTTTGGATCTTTTTCAGGCAACACCTGGACTAT 600
541 GCAGTCTTAATGCTCTCTTGGGAAGCCTCTTTTGGATCTTTTTCAGGCAACACCTGGACTAT 600
601 GGAATGTAATCTCGAGAAGCAACTATTAAGAGAAAGAAAGCAATTTGGAACAGTCGGA 660
601 GGAATGTAATCTCGAGAAGCAACTATTAAGAGAAAGAAAGCAATTTGGAACAGTCGGA 660
661 ATTGCTCTTTACGATTTATCACCAGGAAGTGGAAACATTTCTGTGTTTCAAGCCGTAAGTGA 720
661 ATTGCTCTTTACGATTTATCACCAGGAAGTGGAAACATTTCTGTGTTTCAAGCCGTAAGTGA 720
721 ATTATCAAGTAAAGATGGAGGCGCAAGAGATTAGCGAAACCTTTAAAGGCCCAAT 780
721 ATTATCAAGTAAAGATGGAGGCGCAAGAGATTAGCGAAACCTTTAAAGGCCCAAT 780
781 CTAGTGGAACTAGTTGTCCTCAACATACGGATGGATCCAAATTTATGCCCTGCTGATCCA 840
781 CTAGTGGAACTAGTTGTCCTCAACATACGGATGGATCCAAATTTATGCCCTGCTGATCCA 840
841 GACTGGATGCTTTTATACATAGCAACGATATTTGGATATCTAACATCGTAACACAGAA 900
841 GACTGGATGCTTTTATACATAGCAACGATATTTGGATATCTAACATCGTAACACAGAA 900
901 GAAAGGAGACTCATTTATGTCACATGAGCTAGCAACATGGAAGAGATGCCAGATCA 960
901 GAAAGGAGACTCATTTATGTCACATGAGCTAGCAACATGGAAGAGATGCCAGATCA 960
961 GCTGGAGTGGTACCTTTGTTCTCCAGAGAAATTTGATAGATTTCTGGCTATTTGGTGG 1020
961 GCTGGAGTGGTACCTTTGTTCTCCAGAGAAATTTGATAGATTTCTGGCTATTTGGTGG 1020
1021 TGTCACAAAGCTCAAAACAATCCCAAGTGGTGAATAATCTTAGAATTTCTATATGAAGAA 1080
1021 TGTCACAAAGCTCAAAACAATCCCAAGTGGTGAATAATCTTAGAATTTCTATATGAAGAA 1080
1081 AATGATGAATCTGAGGTGGAAATTAATCATGTTACATCCCTATGTTGGAAACAAAGGAG 1140
1081 AATGATGAATCTGAGGTGGAAATTAATCATGTTACATCCCTATGTTGGAAACAAAGGAG 1140
1141 GCAGATTCATTCGGTTATCTGTAAGAGATCATAGATGTCATAGATTAAGAACTAAT 1200
1141 GCAGATTCATTCGGTTATCTTAAACAGGTACAGCAAAATCCTTAAAGTCACTTTTAAGATG 1200
1201 TCAGAAATATGATGATGCTGAAGGAGTGAATATATGTCAGAGCTGGATGGACTCCT 1260
1201 TCAGAAATATGATGATGCTGAAGGAGTGAATATATGTCAGAGCTGGATGGACTCCT 1260
1261 CAACCTTTTGAGATTTCTATTTGAAGGAGTTGAATATATGTCAGAGCTGGATGGACTCCT 1320
1261 CAACCTTTTGAGATTTCTATTTGAAGGAGTTGAATATATGTCAGAGCTGGATGGACTCCT 1320
1321 GAGGAAATATGCTTGGTCCATCTTACATAGATGCTGCGTCCAGACTCGGCTACAGATAGTG 1380
1321 GAGGAAATATGCTTGGTCCATCTTACATAGATGCTGCGTCCAGACTCGGCTACAGATAGTG 1380
1381 TTGATCTCACTGAATTTATATCCAGTAGAAGATGATGTTATGGAAGGCGAGAGACTC 1440
1381 TTGATCTCACTGAATTTATATCCAGTAGAAGATGATGTTATGGAAGGCGAGAGACTC 1440
1441 ATTGAGTCAGTGGCTGATTTGTCAGGCACTAAATTTATCTATGGAAGAAACACAGACATC 1500
1441 ATTGAGTCAGTGGCTGATTTGTCAGGCACTAAATTTATCTATGGAAGAAACACAGACATC 1500
1501 TGGATAAATATCCATGACATCTTTTCAATGTTTTCCTCCAAAGTCAGAAAGAGAAATTTGAG 1560
1501 TGGATAAATATCCATGACATCTTTTCAATGTTTTCCTCCAAAGTCAGAAAGAGAAATTTGAG 1560
1561 TTTATTTTGGCTCTGAATGCAAAACAGGTTTCCGTCATTTATACAAATTTACATCTATT 1620
1561 TTTATTTTGGCTCTGAATGCAAAACAGGTTTCCGTCATTTATACAAATTTACATCTATT 1620
1621 TTAAAGGAAAGCAATATAACGATCCAGTGGTGGGCTGCTCCCAAGTGATTTCAAG 1680

1621 TTAAAGGAAAGCAATATAACGATCCAGTGGTGGGCTGCTGCTCCAAGTGATTTCAAG 1680
1681 TGTCTTATCAAGAGAGAGATAGCAATTAACAGTGGTGAATGGGAAGTTCTTGGCCGGCAT 1740
1681 TGTCTTATCAAGAGAGAGATAGCAATTAACAGTGGTGAATGGGAAGTTCTTGGCCGGCAT 1740
1741 GGATCTAATATCAAGTGGTGAATGCAAGGCTGGTATATTTTGAAGGCAACCAAGAC 1800
1741 GGATCTAATATCAAGTGGTGAATGCAAGGCTGGTATATTTTGAAGGCAACCAAGAC 1800
1801 TCCCTTTTAGAGCATCACCTGTACGTAGTTCAGTTACGTAAATCTCGAGAGGTGACAAGG 1860
1801 TCCCTTTTAGAGCATCACCTGTACGTAGTTCAGTTACGTAAATCTCGAGAGGTGACAAGG 1860
1861 CTGACTGACCGTGGCTACTCACATTTCTTGTGCTGCATCAGTCAGCACTGCTCTTTATA 1920
1861 CTGACTGACCGTGGCTACTCACATTTCTTGTGCTGCATCAGTCAGCACTGCTCTTTATA 1920
1921 AGTAAGTATAGTAACCAAGAAATCCACACTGTGTGCTCCCTTTTCAAGCTATCAAGTCT 1980
1921 AGTAAGTATAGTAACCAAGAAATCCACACTGTGTGCTCCCTTTTCAAGCTATCAAGTCT 1980
1981 GAAGATGACCCCAACTTTGCAAAACAAAGGAATTTTGGGCCACCATTTTGGATTCAGCAGG 2040
1981 GAAGATGACCCCAACTTTGCAAAACAAAGGAATTTTGGGCCACCATTTTGGATTCAGCAGG 2040
2041 CCTCTTCTGACTATATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTTGGATTTACA 2100
2041 CCTCTTCTGACTATATCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTTGGATTTACA 2100
2101 TTGTATGGAGATGCTCTAAGGCTCATGATCTAAGCCTGGAAGAAATATCTTACTGTG 2160
2101 TTGTATGGAGATGCTCTAAGGCTCATGATCTAAGCCTGGAAGAAATATCTTACTGTG 2160
2161 CTGTTCAATATGCTGGTCTCTCAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2218
2161 CTGTTCAATATGCTGGTCTCTCAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2218
2219 AGTATTTCCGCTTGAATACCTAGCTCTCTAGGTTATGCTGGTGTAGTGATAGACAACA 2278
2221 AGTATTTCCGCTTGAATACCTAGCTCTCTAGGTTATGCTGGTGTAGTGATAGACAACA 2280
2279 GGGATCCTCTACCCGAGGCTTAAATTTGAAGGCGCTTTAAATATATAAATGGGTCAA 2338
2281 GGGATCCTCTACCCGAGGCTTAAATTTGAAGGCGCTTTAAATATATAAATGGGTCAA 2340
2339 TAGAAATTTGAGATCAGGTGGAAGGCTCCCAATATCTAGCTTCTCGATATCATTTCAATG 2398
2341 TAGAAATTTGAGATCAGGTGGAAGGCTCCCAATATCTAGCTTCTCGATATCATTTCAATG 2400
2399 ACTTAGATCTGTGGGCTCCACGGCTGCTTATGGAGGATACCTCTCCCTGATGGCAT 2458
2401 ACTTAGATCTGTGGGCTCCACGGCTGCTTATGGAGGATACCTCTCCCTGATGGCAT 2460
2459 TAATGACAGAGTCAAGATATCTTCAAGGTTGCTTATGCTGGGCGCCCACTCTCTGGA 2518
2461 TAATGACAGAGTCAAGATATCTTCAAGGTTGCTTATGCTGGGCGCCCACTCTCTGGA 2520
2519 TCTTCTATGATACAGGATACAGGACGTTATATGGTTCACCTGACCAAGATGAACAGG 2578
2521 TCTTCTATGATACAGGATACAGGACGTTATATGGTTCACCTGACCAAGATGAACAGG 2580
2579 GCTATTACTTTAGGATCTGTGGCCATCCAAAGCAGAAAAAGTTCCCTCTGAAACCAATCGTT 2638
2581 GCTATTACTTTAGGATCTGTGGCCATCCAAAGCAGAAAAAGTTCCCTCTGAAACCAATCGTT 2640
2639 TACTGCTCTTACATGGTTTCTGGAAGAAATGTCATTTTGGCAATACCAAGTATTTAC 2698
2641 TACTGCTCTTACATGGTTTCTGGAAGAAATGTCATTTTGGCAATACCAAGTATTTAC 2700
2699 TGAGTTTTTATGAGGCTGGAAAGCCATATGATTTACAGATCTATCTCCTCAGGAGAC 2758

Db 2701 TGAGTTTTTTAGTGAGGCTCGAAAGCCATATGATTTACAGATCTATCTCTCAGGAGAC 2760
Qy 2759 ACAGCATAGAGTTCTCGATTCGGGAGAACATTAAGACTGCACTCTTTGCACTACCTTC 2818
Db 2761 ACAGCATAGAGTTCTCGAATCGGGAGAACATTAAGACTGCACTCTTTGCACTACCTTC 2820
Qy 2819 AAGAAACCTTTGGATCAGTATTCGCTCTCTAAAGTGATATAATTTGACCTGTGTAGA 2878
Db 2821 AAGAAACCTTTGGATCAGTATTCGCTCTCTAAAGTGATATAATTTGACCTGTGTAGA 2880
Qy 2879 ACTCTCTGTATACACTGGCTATTTAAACCAATGAGGAGTTTAATCAACAGAAAAACA 2938
Db 2881 ACTCTCTGTATACACTGGCTATTTAAACCAATGAGGAGTTTAATCAACAGAAAAACA 2940
Qy 2939 GAATTGATCATCACATTTTGATACCTGCCATGTAAACATCTACTCTCTGAAAAATAAATGTGG 2998
Db 2941 GAATTGATCATCACATTTTGATACCTGCCATGTAAACATCTACTCTCTGAAAAATAAATGTGG 3000
Qy 2999 TGCCATGCAGGGTCTACGGTTTGTGGTAGTAATCTAATACCTTTAAACCCCAATGCTCAA 3058
Db 3001 TGCCATGCAGGGTCTACGGTTTGTGGTAGTAATCTAATACCTTTAAACCCCAATGCTCAA 3060
Qy 3059 ATCAATGATACATATTCTCGAGAGACCCAGCAATACCAATGAATTTACTTAAAAAATA 3118
Db 3061 ATCAATGATACATATTCTCGAGAGACCCAGCAATACCAATGAATTTACTTAAAAAATA 3120
Qy 3119 AA 3120
Db 3121 AA 3122

RESULT 5

US-10-170-789-37
; Sequence 37, Application US/10170789
; Publication No. US20030180930A1
; GENERAL INFORMATION:
; APPLICANT: Rachel E. Meyers
; APPLICANT: Olandt, Peter J.
; APPLICANT: Kapeller-Liebermann, Rosana
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Williamson, Mark
; APPLICANT: Weich, Nadine
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASE, PHOSPHATASE,
; AND PROTEASE FAMILY MEMBERS AND USES THEREOF
; FILE REFERENCE: 10448-191001
; CURRENT APPLICATION NUMBER: US/10/170,789
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: US 09/797,039
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: PCT/US01/06525
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/186,061
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 09/882,166
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: PCT/US01/19269
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/212,078
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/934,406
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: PCT/US01/26052
; PRIOR FILING DATE: 2000-08-21
; PRIOR APPLICATION NUMBER: US 60/226,740
; PRIOR FILING DATE: 2000-08-21
; PRIOR APPLICATION NUMBER: US 09/861,801
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: PCT/US01/16549
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: US 60/205,508
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 09/801,267
; PRIOR FILING DATE: 2001-03-06

; PRIOR APPLICATION NUMBER: PCT/US01/07138
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: US 60/187,454
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: US 09/829,671
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: PCT/US01/40483
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: US 60/197,508
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: US 09/961,721
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: PCT/US01/29904
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: US 60/235,023
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US 10/045,367
; PRIOR FILING DATE: 2001-11-07
; PRIOR APPLICATION NUMBER: US 60/246,561
; PRIOR FILING DATE: 2000-11-07
; PRIOR APPLICATION NUMBER: US 09/801,275
; PRIOR FILING DATE: 2001-03-06
; PRIOR APPLICATION NUMBER: PCT/US01/07074
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: US 60/187,420
; PRIOR FILING DATE: 2000-03-07
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 37
; LENGTH: 3143
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (229) ... (2874)
US-10-170-789-37

Query Match 99.5%; Score 3105.2; DB 6; Length 3143;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3107; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 11 GCCTCCGAGCCCAAGCCGCTCTACTGCGCGCGCTCTTCTTAGTGCCCGTTCCCGC 70
Db 26 GCCTCCGAGCCCAAGCCGCTCTACTGCGCGCGCTCTTCTTAGTGCCCGTTCCCGC 85

Qy 71 CTGGGTTGTCACCGCGCGCGCGAGGAGCCACTGCAACCGACGCGAGTGGAGGC 130
Db 86 CTGGGTTGTCACCGCGCGCGCGAGGAGCCACTACACCGAGCCGAGTGGAGGC 145

Qy 131 GCGCAGCATGAAGCGCGCGCGAGGCGCTCCATAGCGCAGTCCGAGACGCTCCGCGCGG 190
Db 146 GCGCAGCATGAAGCGCGCGCGAGGCGCTCCATAGCGCAGTCCGAGACGCTCCGCGCGG 205

Qy 191 GCGGGGGGAAGGAAAAATGCAACATGGCAGCAATGGAAACAGAAACAGCTGGGTGTG 230
Db 206 GCGGGGGGAAGGAAAAATGCAACATGGCAGCAATGGAAACAGAAACAGCTGGGTGTG 265

Qy 251 AGATATTTGAAACTCGGACCTGTGAGGAGAAATTTGAATCAAGGATCGCGCTAAATTTG 310
Db 266 AGATATTTGAAACTCGGACCTGTGAGGAGAAATTTGAATCAAGGATCGCGCTAAATTTG 325

Qy 311 AGCTTTTATTTGTTGAGCGGTATTCCTGGAGTCAAGCTTAAAAAGCTCTTGGCGATACCA 370
Db 326 AGCTTTTATTTGTTGAGCGGTATTCCTGGAGTCAAGCTTAAAAAGCTCTTGGCGATACCA 385

Qy 371 GAAATATCATGTGCTACATGATGGCTAAGGACCAACATGATTTTCATTTGTGTAAGAGGA 430
Db 386 GAAATATCATGTGCTACATGATGGCTAAGGACCAACATGATTTTCATTTGTGTAAGAGGA 445

Qy 431 ATGATCCAGATGGACCTCATTCCAGACAGAATCTATTACCTTGGCCATCTCTGGTGAGAA 490
Db 446 ATGATCCAGATGGACCTCATTCCAGACAGAATCTATTACCTTGGCCATCTCTGGTGAGAA 505

Qy 491 GAGAAATACACGTGTTTTATTCTGAAATCCCAAAATCCATATAGAGAGAGAGCTTTAA 550
Db 506 GAGAAATACACGTGTTTTATTCTGAAATCCCAAAATCCATATAGAGAGAGAGCTTTAA 565
Qy 551 TGCTCTCTTGGAGCCCTCTTTTGGATCTTTTTCAGGCAACACTGCACTATGGAATGTATT 610
Db 566 TGCTCTCTTGGAGCCCTCTTTTGGATCTTTTTCAGGCAACACTGCACTATGGAATGTATT 625
Qy 611 CTCGAGAAGAAGAACTATTAAAGAGAAGAAACGATTTGGAAACAGTCTGGAAATTCCTTT 670
Db 626 CTCGAGAAGAAGAACTATTAAAGAGAAGAAACGATTTGGAAACAGTCTGGAAATTCCTTT 685
Qy 671 ACGATTTACACGAAGAAGTGAACATTTCTGTGTTCAAGCCGGTAGTGGAAATTTATCAG 730
Db 686 ACGATTTATCACGAAGAAGTGAACATTTCTGTGTTCAAGCCGGTAGTGGAAATTTATCAG 745
Qy 731 TAAAGATGGAGGCCACAAAGGATTTAGCAACACCTTTTAAAGCCCAATCTAGTGGAA 790
Db 746 TAAAGATGGAGGCCACAAAGGATTTAGCAACACCTTTTAAAGCCCAATCTAGTGGAA 805
Qy 791 CTAGTTGTCCCAACATACCGATGGATCCAAATTTATGCCCGCTGTATCCAGACTGGATTG 850
Db 806 CTAGTTGTCCCAACATACCGATGGATCCAAATTTATGCCCGCTGTATCCAGACTGGATTG 865
Qy 851 CTTTTTATACATAGCAACGATATTTGGATATCTAACATCGTAAACAGAGAGAAAGGAGAC 910
Db 866 CTTTTTATACATAGCAACGATATTTGGATATCTAACATCGTAAACAGAGAGAAAGGAGAC 925
Qy 911 TCACTTATGTGCAATAGCTAGCCCAATGGAGAGATGCCAGATCAGCTGGAGTCG 970
Db 926 TCACTTATGTGCAATAGCTAGCCCAATGGAGAGATGCCAGATCAGCTGGAGTCG 985
Qy 971 CTACCTTTGTTCTCCAAGAGAAATTTGATAGATATTTGGGCTATTTGGTGTGTCCTCAAAAG 1030
Db 986 CTACCTTTGTTCTCCAAGAGAAATTTGATAGATATTTGGGCTATTTGGTGTGTCCTCAAAAG 1045
Qy 1031 CTGAAACAACTCCACGTGGTGTGTAATTTCTTGAATTTCTATATGAAGAAATATGATGAAT 1090
Db 1046 CTGAAACAACTCCACGTGGTGTGTAATTTCTTGAATTTCTATATGAAGAAATATGATGAAT 1105
Qy 1091 CTGAGGTGGAAATTTATGTTTACATCCCTATGTTGGAAACAAAGAGGCGAGATTCAT 1150
Db 1106 CTGAGGTGGAAATTTATGTTTACATCCCTATGTTGGAAACAAAGAGGCGAGATTCAT 1165
Qy 1151 TCCGTTATCTTAAACAGGTACAGCAATCTTAAAGTCACTTTTAAAGATGTCAAGAAATA 1210
Db 1166 TCCGTTATCTTAAACAGGTACAGCAATCTTAAAGTCACTTTTAAAGATGTCAAGAAATA 1225
Qy 1211 TGATTTGATGCTGAAGGAAGGATCATAGATGTCATAGATAGAGAACTAAATTCACCTTTTG 1270
Db 1226 TGATTTGATGCTGAAGGAAGGATCATAGATGTCATAGATAGAGAACTAAATTCACCTTTTG 1285
Qy 1271 AGATTTCTATTTGAAGAGTTGAATATATTTGCCAGAGCTGGATGGAATCTCTGAGGGAAAT 1330
Db 1286 AGATTTCTATTTGAAGAGTTGAATATATTTGCCAGAGCTGGATGGAATCTCTGAGGGAAAT 1345
Qy 1331 ATGCTTGGTCCATCTTACTAGATCGCTCCAGCTCGCTACAGATAGTGTGATCTCAC 1390
Db 1346 ATGCTTGGTCCATCTTACTAGATCGCTCCAGCTCGCTACAGATAGTGTGATCTCAC 1405
Qy 1391 CTGAATTTATTTATCCAGTAGAAGATGATGTTATGGAAGGCGAGAGACTCAATGAGTCAG 1450
Db 1406 CTGAATTTATTTATCCAGTAGAAGATGATGTTATGGAAGGCGAGAGACTCAATGAGTCAG 1465
Qy 1451 TGCTGTATCTGTGAGCCCACTAATTTATCTATGAAGAAACAAAGACATCTGGATAAATA 1510
Db 1466 TGCTGTATCTGTGAGCCCACTAATTTATCTATGAAGAAACAAAGACATCTGGATAAATA 1525
Qy 1511 TCCATGACATCTTTCAATGTTTTTCCCAAGTCAGAGAGAAATTTGAGTTATTTTTTG 1570
Db 1526 TCCATGACATCTTTCAATGTTTTTCCCAAGTCAGAGAGAAATTTGAGTTATTTTTTG 1585
Qy 1571 CCTCTGAATGCAAAACAGGTTTCCGTCATTTTATACAAATTTACATCTATTTTAAAGGAAA 1630

Db 1586 CCTCTGAATGCAAAACAGGTTTCCGTCATTTTATACAAATTTACATCTATTTTAAAGGAAA 1645
Qy 1631 GCAAAATATAAACGATCCAGTGGTGGGCTCCCTGCTCCAAAGTGATTTCAAGTGTCTTATCA 1690
Db 1646 GCAAAATATAAACGATCCAGTGGTGGGCTCCCTGCTCCAAAGTGATTTCAAGTGTCTTATCA 1705
Qy 1691 AAGAGAGATAGCAATTTACCAGTGGTGAATGGGAAGTCTTTGGCGGCGCATGGATCTAATA 1750
Db 1706 AAGAGAGATAGCAATTTACCAGTGGTGAATGGGAAGTCTTTGGCGGCGCATGGATCTAATA 1765
Qy 1751 TCCAAGTTGATGAAGTCAGAAGGCTGGTATATTTTGAAGGCAACCAAGAGCTCCCTCTTTAG 1810
Db 1766 TCCAAGTTGATGAAGTCAGAAGGCTGGTATATTTTGAAGGCAACCAAGAGCTCCCTCTTTAG 1825
Qy 1811 AGCATCACCTGTACGTAGTCAGTTACGTAAATTCCTGGAGAGTGACAAAGCGCTGACTGACC 1870
Db 1826 AGCATCACCTGTAGTCAGTTACGTAAATTCCTGGAGAGTGACAAAGCGCTGACTGACC 1885
Qy 1871 GTGGCTATCTACATTTCTTCTGCTGATCAGTCAGCACTGTGACTTCTTTTATTAAGTAAGTATA 1930
Db 1886 GTGGCTACTCACATTTCTTCTGCTGATCAGTCAGCACTGTGACTTCTTTTATTAAGTAAGTATA 1945
Qy 1931 GTAAACAGAGAGATCCACACTGTGTGCCCTTTTAAAGCTATCAAGCTCCCTGGAAGTACCC 1990
Db 1946 GTAAACAGAGAGATCCACACTGTGTGCCCTTTTAAAGCTATCAAGCTCCCTGGAAGTACCC 2005
Qy 1991 CAACTTGCACAAAACAAAGGAATTTTGGGCCACCATTTTGGATTCAGCAGGTCCTCTTCTCTG 2050
Db 2006 CAACTTGCACAAAACAAAGGAATTTTGGGCCACCATTTTGGATTCAGCAGGTCCTCTTCTCTG 2065
Qy 2051 ACTATACTCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGAATTTACATGTTATGGGA 2110
Db 2066 ACTATACTCTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGAATTTACATGTTATGGGA 2125
Qy 2111 TGCTCTACAGCCCTCATGATCTACAGCTTGGAAAGAAATATCCCTGCTGTGTTCTATAT 2170
Db 2126 TGCTCTACAGCCCTCATGATCTACAGCTTGGAAAGAAATATCCCTGCTGTGTTCTATAT 2185
Qy 2171 ATGCTGTCTCTCAGGTGCAAGTTGGTGAATAATTCGGTTTAAAGGAGTCAAGTATTTCCGCT 2230
Db 2186 ATGCTGTCTCTCAGGTGCAAGTTGGTGAATAATTCGGTTTAAAGGAGTCAAGTATTTCCGCT 2245
Qy 2231 TGAATACCTCTAGCTCTCTAGGTATATGTTGTTAGTATAGACAAACAGGGGATCTGTCTC 2290
Db 2246 TGAATACCTCTAGCTCTCTAGGTATATGTTGTTAGTATAGACAAACAGGGGATCTGTCTC 2305
Qy 2291 ACCGAGGCTTTAAATTTGAAGGCGCTTTTAAATATAAAATGGGTCAATAGAAATTCAGC 2350
Db 2306 ACCGAGGCTTTAAATTTGAAGGCGCTTTTAAATATAAAATGGGTCAATAGAAATTCAGC 2365
Qy 2351 ATCAGGTGGAAGGACTCCCAATATCTAGCTTCTCGATATGATTTTCAATGACTTAGATCGTG 2410
Db 2366 ATCAGGTGGAAGGACTCCCAATATCTAGCTTCTCGATATGATTTTCAATGACTTAGATCGTG 2425
Qy 2411 TGGGCATCCACGCTGCTCTATGGAAGTACCTCTCCCTGATGGCATTAATGACAGAGGT 2470
Db 2426 TGGGCATCCACGCTGCTCTATGGAAGTACCTCTCCCTGATGGCATTAATGACAGAGGT 2485
Qy 2471 CAGATATCTTTCAGGGTGTCTATGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATA 2530
Db 2486 CAGATATCTTTCAGGGTGTCTATGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATA 2545
Qy 2531 CAGGATACAGGAACTGATATGGGTCACTCCCTGACCAAGTGAACAGGGCTTACTTTAG 2590
Db 2546 CAGGATACAGGAACTGATATGGGTCACTCCCTGACCAAGTGAACAGGGCTTACTTTAG 2605
Qy 2591 GATCTGGGCCATGCAAGCAGAAAAGTTCCTCTCTGAAACCAAAATGCTTTTACTGCTTTAC 2650
Db 2606 GATCTGGGCCATGCAAGCAGAAAAGTTCCTCTCTGAAACCAAAATGCTTTTACTGCTTTAC 2665
Qy 2651 ATGCTTCTCTGGATGAGAAATGTCATTTTGGCATCATCAAGTATATTACTGAGTTTCTTTAG 2710

Db 902 ACTTATGTGCAAAATGAGCTAGCCAAATCGAAGAAGATGCCAGATCAGCTGGAGTCGCT 961
Qy 973 ACCTTTGTGTTCCCAAGAAGAAATTTGATAGATATTTCTGGCTATTGTTGGTGTCCAAAGCT 1032
Db 962 ACCTTTGTGTTCCCAAGAAGAAATTTGATAGATATTTCTGGCTATTGTTGGTGTCCAAAGCT 1021
Qy 1033 GAAACAACTCCAGCTGGTGGTAAATTTCTTAGAATTTCTATATGAAGAAATGATGAATCT 1092
Db 1022 GAAACAACTCCAGCTGGTGGTAAATTTCTTAGAATTTCTATATGAAGAAATGATGAATCT 1081
Qy 1093 GAGGTGGAATTTATTTCAATGTTACATCCCTATGTTGGAACAAGGAGGCGAGATTCATTC 1152
Db 1082 GAGGTGGAATTTATTTCAATGTTACATCCCTATGTTGGAACAAGGAGGCGAGATTCATTC 1141
Qy 1153 CGTTATCCTAAAAACAGGTACAGCAAAATCCTAAAGTCACTTTTAAAGTGTGAGAAATAATG 1212
Db 1142 CGTTATCCTAAAAACAGGTACAGCAAAATCCTAAAGTCACTTTTAAAGTGTGAGAAATAATG 1201
Qy 1213 ATTTGATGTGAAGAGGATCATAGATGTGATAGATAGAGAACTAAATTCAACTTTTGGAG 1272
Db 1202 ATTTGATGTGAAGAGGATCATAGATGTGATAGATAGAGAACTAAATTCAACTTTTGGAG 1261
Qy 1273 ATTTCTATTTGAAGGTTGAATATATTTGCCAGAGCTGGATGGACTCTCGAGGGAATAAT 1332
Db 1262 ATTTCTATTTGAAGGTTGAATATATTTGCCAGAGCTGGATGGACTCTCGAGGGAATAAT 1321
Qy 1333 GCTTGGTCCATCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCT 1392
Db 1322 GCTTGGTCCATCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACCT 1381
Qy 1393 GAAATTTATTTATCCCAAGTGAAGATGATGTTATGGAAGGAGAGATCTCATTTGAGTCAGTG 1452
Db 1382 GAAATTTATTTATCCCAAGTGAAGATGATGTTATGGAAGGAGAGATCTCATTTGAGTCAGTG 1441
Qy 1453 CTGTATCTGTGAGCCCACTAAATTTATCTATGAAGAAACAAGATCTGATTAATATC 1512
Db 1442 CCGTATCTGTGAGCCCACTAAATTTATCTATGAAGAAACAAGATCTGATTAATATC 1501
Qy 1513 CATGACATCTTTCAATGTTTCCCAAGTCAAGAGGAAATTCAGTTATTTTGGC 1572
Db 1502 CATGACATCTTTCAATGTTTCCCAAGTCAAGAGGAAATTCAGTTATTTTGGC 1561
Qy 1573 TCTGAATCAAAACAGGTTTCCGTCATTTATACAAAATTACATCTATTTTAAAGGAAAGC 1632
Db 1562 TCTGAATCAAAACAGGTTTCCGTCATTTATACAAAATTACATCTATTTTAAAGGAAAGC 1621
Qy 1633 AAATATAAACGATCCAGTGGTGGCTGCTCCAAAGTGAATTTCAAGTGTCTCTATCAAA 1692
Db 1622 AAATATAAACGATCCAGTGGTGGCTGCTCCAAAGTGAATTTCAAGTGTCTCTATCAAA 1681
Qy 1693 GAGGAGATAGCAATTTACAGTGGTGAATGGGAAGTCTTGGCCGCATGGATCTAATATC 1752
Db 1682 GAGGAGATAGCAATTTACAGTGGTGGTGAATGGGAAGTCTTGGCCGCATGGATCTAATATC 1741
Qy 1753 CAAGTTGATGAAGTCAAGAGGCTGGTATATTTTGAAGGCCAAAGACTCCCTTTTGAAG 1812
Db 1742 CAAGTTGATGAAGTCAAGAGGCTGGTATATTTTGAAGGCCAAAGACTCCCTTTTGAAG 1801
Qy 1813 CATCACCTGTAGTACGTTACGTTAAATCCTCGAGAGGTGAACAAGGCTGACTGACCGT 1872
Db 1802 CATCACCTGTAGTACGTTACGTTAAATCCTCGAGAGGTGAACAAGGCTGACTGACCGT 1861
Qy 1873 GGTACTCATCTTGTGTCATAGTCAGTCAGCACGTGCACTTCTTATTAAGTAGTATAGT 1932
Db 1862 GGTACTCATCTTGTGTCATAGTCAGTCAGCACGTGCACTTCTTATTAAGTAGTATAGT 1921
Qy 1933 AACCGAAGATCCACACTGTGTGTCCTTTTACAGCTATCAAGCTCTGAAGATGACCCA 1992
Db 1922 AACCGAAGATCCACACTGTGTGTCCTTTTACAGCTATCAAGCTCTGAAGATGACCCA 1981
Qy 1993 ACTTGCAGAAACAAAGAAATTTTGGGCCACCACTTTTGGATTTCAGCAGGTCTCTTCTGAC 2052
Db 1982 ACTTGCAGAAACAAAGAAATTTTGGGCCACCACTTTTGGATTTCAGCAGGTCTCTTCTGAC 2041

Qy 2053 TATACTCTCCAGAAATTTTCTCTTTTGAAAGTACTACTGGATTTACATTTGATGGGATG 2112
Db 2042 TATACTCTCCAGAAATTTTCTCTTTTGAAAGTACTACTGGATTTACATTTGATGGGATG 2101
Qy 2113 CTCTACAAGCCTCATGATCTACAGCCTGGAAGAAATATCTCTACTGTGCTGTTTCATATAT 2172
Db 2102 CTCTACAAGCCTCATGATCTACAGCCTGGAAGAAATATCTCTACTGTGCTGTTTCATATAT 2161
Qy 2173 GGTGTGCTTCAGGTGCAGTGGTGAATTAATCGGTTTAAAGAGTCAAGTATTTCCGCTTG 2232
Db 2162 GGTGTGCTTCAGGTGCAGTGGTGAATTAATCGGTTTAAAGAGTCAAGTATTTCCGCTTG 2221
Qy 2233 AATACCTTAGCCTCTCTAGTGTATGTGTTGTAGTGTAGATAGACACAGGGGATCTGTGCAC 2292
Db 2222 AATACCTTAGCCTCTCTAGTGTATGTGTTGTAGTGTAGATAGACACAGGGGATCTGTGCAC 2281
Qy 2293 CGAGGGCTTTAAATTTGAAAGGCGCTTTAAATATATAAATGGGTCAAAATAGAAAATTGACGAT 2352
Db 2282 CGAGGGCTTTAAATTTGAAAGGCGCTTTAAATATATAAATGGGTCAAAATAGAAAATTGACGAT 2341
Qy 2353 CAGGTGGAAGGACTCCAAATATCTAGCTTCTCGATATGATTTCAATTGACTTAGATCGTGTG 2412
Db 2342 CAGGTGGAAGGACTCCAAATATCTAGCTTCTCGATATGATTTCAATTGACTTAGATCGTGTG 2401
Qy 2413 GGCATCCAGCGCTGTCCTATGAGGATACCTCTCCCTGATGGCATTAATGTCAGAGGTCA 2472
Db 2402 GGCATCCAGCGCTGTCCTATGAGGATACCTCTCCCTGATGGCATTAATGTCAGAGGTCA 2461
Qy 2473 GATATCTTCAGGGTGTCTATTTGCTGGGCGCCAGTCACTCTGTGGATCTCTTATGATACA 2532
Db 2462 GATATCTTCAGGGTGTCTATTTGCTGGGCGCCAGTCACTCTGTGGATCTCTTATGATACA 2521
Qy 2533 GGATACACGGAACTTTATATGGGTCACTCTGACCAAGATGAACAGGGCTATTACTTAGGA 2592
Db 2522 GGATACACGGAACTTTATATGGGTCACTCTGACCAAGATGAACAGGGCTATTACTTAGGA 2581
Qy 2593 TCTGTGGCCATGCAAGCAAGAAAGTTCCCTCTGAAACCAATTCGTTTACTGCTCTTACAT 2652
Db 2582 TCTGTGGCCATGCAAGCAAGAAAGTTCCCTCTGAAACCAATTCGTTTACTGCTCTTACAT 2641
Qy 2653 GGTTCCTGGATGAGATGTCCATTTTGCACATACAGTATATTACTGAGTTTTTTAGTG 2712
Db 2642 GGTTCCTGGATGAGATGTCCATTTTGCACATACAGTATATTACTGAGTTTTTTAGTG 2701
Qy 2713 AGGCTGGAAGCCATATGATTTTACAGATCTATCTCAGGAGAGACACAGCATAAAGATT 2772
Db 2702 AGGCTGGAAGCCATATGATTTTACAGATCTATCTCAGGAGAGACACAGCATAAAGATT 2761
Qy 2773 CCTGAATTCGGAGAACATTTATGAACCTGTCATCTTTTGCACCTACCTTCAAGAAAACCTTGGGA 2832
Db 2762 CCTGAATTCGGAGAACATTTATGAACCTGTCATCTTTTGCACCTACCTTCAAGAAAACCTTGGGA 2821
Qy 2833 TCAAGTATTTGCTGCTCTTAAAGATGATTAATTTTGCACCTGTGTAGAACCTCTCTGGTATAC 2892
Db 2822 TCAAGTATTTGCTGCTCTTAAAGATGATTAATTTTGCACCTGTGTAGAACCTCTCTGGTATAC 2881
Qy 2893 ACTGGCTATTTAAACCAAAATGAGGAGGTTTAAATCAACAGAAAACACAGAAATTTGATCATCAC 2952
Db 2882 ACTGGCTATTTAAACCAAAATGAGGAGGTTTAAATCAACAGAAAACACAGAAATTTGATCATCAC 2941
Qy 2953 ATTTTGATACCTGCCATGTAAACATCTACTCTCTGAAAAATAAATGTGGTGCCATGACGGGT 3012
Db 2942 ATTTTGATACCTGCCATGTAAACATCTACTCTCTGAAAAATAAATGTGGTGCCATGACGGGT 3001
Qy 3013 CTACGGTTGTGTAGTAAATCTTAATACCTTAACCCCACTACCTGCTCAAAATCAAAATGATACA 3072
Db 3002 CTACGGTTGTGTAGTAAATCTTAATACCTTAACCCCACTACCTGCTCAAAATCAAAATGATACA 3061
Qy 3073 TATTTCTGAGAGACCCAGCAATACCAATAGAAATTTACTAAAAA 3117
Db 3062 TATTTCTGAGAGACCCAGCAATACCAATAGAAATTTACTAAAAA 3106

RESULT 7	US-09-976-674-22	Sequence 22; Application US/09976674	661	ATTGCTTCTTACGATTATCACCAAGGAAGTGAACATTTCTGTCTTCAAGCCGGTAGTGG	720
US-09-976-674-22	Patent No. US20020115643A1	Sequence 22; Application US/09976674	721	ATTTATCACGTAAAGATGAGGGGCCCAAGGATTTTACGCAACAACCTTTTAAAGGCCCAAT	780
US-09-976-674-22	GENERAL INFORMATION:	Sequence 22; Application US/09976674	721	ATTTATCACGTAAAGATGAGGGGCCCAAGGATTTTACGCAACAACCTTTTAAAGGCCCAAT	780
US-09-976-674-22	APPLICANT: Qi, Steve	Sequence 22; Application US/09976674	781	CTAGTGGAACTAGTGTGTCGCAACATACGGAATGATCCAAATTTATGCCCGCTGATCCA	840
US-09-976-674-22	APPLICANT: Akineanya, Karen	Sequence 22; Application US/09976674	781	CTAGTGGAACTAGTGTGTCGCAACATACGGAATGATCCAAATTTATGCCCGCTGATCCA	840
US-09-976-674-22	APPLICANT: Riviere, Pierre	Sequence 22; Application US/09976674	841	GACTGGATTGCTTTTATACATAGCAACGATTTTGGATATCTACATGCTTAACACGAGAA	900
US-09-976-674-22	APPLICANT: Junien, Jean-Louis	Sequence 22; Application US/09976674	841	GACTGGATTGCTTTTATACATAGCAACGATTTTGGATATCTACATGCTTAACACGAGAA	900
US-09-976-674-22	FILE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV	Sequence 22; Application US/09976674	901	GAAAGGAGACTCATTATGTGCAACATGAGCTAGCCCAACATGGAAGAAAGATGCCAGATCA	960
US-09-976-674-22	FILE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV	Sequence 22; Application US/09976674	901	GAAAGGAGACTCATTATGTGCAACATGAGCTAGCCCAACATGGAAGAAAGATGCCAGATCA	960
US-09-976-674-22	CURRENT APPLICATION NUMBER: US 09/976,674	Sequence 22; Application US/09976674	961	GCTGAGTGCCTGCTTCTTCCAAAGAAATTTGATAGATATCTGGCTATGTGTGG	1020
US-09-976-674-22	CURRENT APPLICATION NUMBER: US 09/976,674	Sequence 22; Application US/09976674	961	GCTGAGTGCCTGCTTCTTCCAAAGAAATTTGATAGATATCTGGCTATGTGTGG	1020
US-09-976-674-22	PRIOR FILING DATE: 2001-10-12	Sequence 22; Application US/09976674	1021	TGTCAAAAGCTGAAACAACTCCCAAGTGGTGGTAAATTTCTAGAAATCTTATATGAAGAA	1080
US-09-976-674-22	PRIOR FILING DATE: 2001-10-12	Sequence 22; Application US/09976674	1021	TGTCAAAAGCTGAAACAACTCCCAAGTGGTGGTAAATTTCTAGAAATCTTATATGAAGAA	1080
US-09-976-674-22	PRIOR FILING DATE: 2000-10-12	Sequence 22; Application US/09976674	1081	AATGATGATCTGAGGTGGAAATTTATTCATGTATCATCCCCCTATGTTGGAAACAAGGAGG	1140
US-09-976-674-22	PRIOR FILING DATE: 2000-10-12	Sequence 22; Application US/09976674	1081	AATGATGATCTGAGGTGGAAATTTATTCATGTATCATCCCCCTATGTTGGAAACAAGGAGG	1140
US-09-976-674-22	NUMBER OF SEQ ID NOS: 61	Sequence 22; Application US/09976674	1141	GCAGATTCATTCCGTTATCTCTAAACACAGGTACAGCAATCTCTAAAGTCACTTTTAAAGATG	1200
US-09-976-674-22	NUMBER OF SEQ ID NOS: 61	Sequence 22; Application US/09976674	1141	GCAGATTCATTCCGTTATCTCTAAACACAGGTACAGCAATCTCTAAAGTCACTTTTAAAGATG	1200
US-09-976-674-22	SOFTWARE: Patent in version 3.1	Sequence 22; Application US/09976674	1201	TCAGAAATAATGATTGATGCTGAAGGAAGGATCATAGATGTCATAGATTAAGGAATCTAATTT	1260
US-09-976-674-22	SOFTWARE: Patent in version 3.1	Sequence 22; Application US/09976674	1201	TCAGAAATAATGATTGATGCTGAAGGAAGGATCATAGATGTCATAGATTAAGGAATCTAATTT	1260
US-09-976-674-22	SEQ ID NO 22	Sequence 22; Application US/09976674	1261	CAACCTTTTGAGATCTTATTTGAAGGAGTTCGAATATATTTGCCAGAGCTGGATGCACTCCT	1320
US-09-976-674-22	SEQ ID NO 22	Sequence 22; Application US/09976674	1261	CAACCTTTTGAGATCTTATTTGAAGGAGTTCGAATATATTTGCCAGAGCTGGATGCACTCCT	1320
US-09-976-674-22	LENGTH: 4685	Sequence 22; Application US/09976674	1321	GAGGGAATAATGCTTTGGTCCATCTCTAGATCGCTCCAGACTCGCTCCAGACTCGCTACAGATAGT	1380
US-09-976-674-22	LENGTH: 4685	Sequence 22; Application US/09976674	1321	GAGGGAATAATGCTTTGGTCCATCTCTAGATCGCTCCAGACTCGCTCCAGACTCGCTACAGATAGT	1380
US-09-976-674-22	TYPE: DNA	Sequence 22; Application US/09976674	1381	TTGATCTCACGTGAATTTATTTCCAGTGAAGATGATGTTATGGAAGGACAGAGACTC	1440
US-09-976-674-22	TYPE: DNA	Sequence 22; Application US/09976674	1381	TTGATCTCACGTGAATTTATTTCCAGTGAAGATGATGTTATGGAAGGACAGAGACTC	1440
US-09-976-674-22	ORGANISM: Homo sapiens	Sequence 22; Application US/09976674	1441	ATTGAGTCAGTCCCTGATTCGTGAGCCACTTAATTTATCTATGAGAAACAACAGACATC	1500
US-09-976-674-22	ORGANISM: Homo sapiens	Sequence 22; Application US/09976674	1441	ATTGAGTCAGTCCCTGATTCGTGAGCCACTTAATTTATCTATGAGAAACAACAGACATC	1500
US-09-976-674-22	Best Local Similarity 90.5%; Score 2824.4; DB 3; Length 4685;	Sequence 22; Application US/09976674	1501	TGATATAATCCATGACATCTTTTCATGTTTTCCTCCCAAGTCCAGAGAGAGAAATTTGAG	1560
US-09-976-674-22	Best Local Similarity 90.5%; Score 2824.4; DB 3; Length 4685;	Sequence 22; Application US/09976674	1501	TGATATAATCCATGACATCTTTTCATGTTTTCCTCCCAAGTCCAGAGAGAGAAATTTGAG	1560
US-09-976-674-22	Mismatches 0; Indels 142; Gaps 1;	Sequence 22; Application US/09976674	1561	TTTATTTTGGCTCTGAAATGCAAAAACAGGTTTCCGTGATTTTATACAAAATTTACATCTATT	1620
US-09-976-674-22	Mismatches 0; Indels 142; Gaps 1;	Sequence 22; Application US/09976674	1561	TTTATTTTGGCTCTGAAATGCAAAAACAGGTTTCCGTGATTTTATACAAAATTTACATCTATT	1620
US-09-976-674-22	Mismatches 0; Indels 142; Gaps 1;	Sequence 22; Application US/09976674	1621	TTAAAGGAAGCAAAATATAAACCGATCCAGTGGTGGGCTGCCTGCTCCAAAGTGAATTTCAAG	1680
US-09-976-674-22	Mismatches 0; Indels 142; Gaps 1;	Sequence 22; Application US/09976674	1621	TTAAAGGAAGCAAAATATAAACCGATCCAGTGGTGGGCTGCCTGCTCCAAAGTGAATTTCAAG	1680
US-09-976-674-22	Mismatches 0; Indels 142; Gaps 1;	Sequence 22; Application US/09976674	1681	TGTCCTTATCAAAAGAGAGATAGCAATTAACCGATGGTGAATGGGAAGTCTTTGGCGGGCAT	1740
US-09-976-674-22	Mismatches 0; Indels 142; Gaps 1;	Sequence 22; Application US/09976674	1681	TGTCCTTATCAAAAGAGAGATAGCAATTAACCGATGGTGAATGGGAAGTCTTTGGCGGGCAT	1740
US-09-976-					

1741	GGATCTAAATATCCAAAGTTGATGAAGTCAGAGCGCTGGTATATATTTTGAAGGCCACCAAGAC	1800
Db		
1801	TCCTCTTTAGAGCATCACCTGTACGTAGTCAGTTAGCTAAATCTCGGAGAGTGACAAAGG	1860
Qy		
1801	TCCTCTTTAGAGCATCACCTGTACGTAGTCAGTTAGCTAAATCTCGGAGAGTGACAAAGG	1860
Db		
1861	CTGACTGACCGTGGCTACTCACAATTCCTGCTGCATCAGTCAGCACTGTGACTTCCTTTTATA	1920
Qy		
1861	CTGACTGACCGTGGCTACTCACAATTCCTGCTGCATCAGTCAGCACTGTGACTTCCTTTTATA	1920
Db		
1921	AGTAAGTATAGTAACCAAGAAATCCACACTGTGTGTCCTTTTACAAGCTATCAAGTCCT	1980
Qy		
1921	AGTAAGTATAGTAACCAAGAAATCCACACTGTGTGTCCTTTTACAAGCTATCAAGTCCT	1980
Db		
1981	GAAGATGACCCAACTTGCACAAACAAAGAAATTTTGGGCCACCAATTTGGATTTCAGCAGGT	2040
Qy		
1981	GAAGATGACCCAACTTGCACAAACAAAGAAATTTTGGGCCACCAATTTGGATTTCAGCAGGT	2040
Db		
2041	CCTCTTCCTGACTATATCTCCTCMAGAAATTTTCTCTTTTGAAGTACTACTGGATTTTACA	2100
Qy		
2035	-----	2034
Db		
2101	TTGTATGGATGCTCTACAGGCTCATGATCTACAGCTCGGAAAGAAATATCCTACTGTG	2160
Qy		
2035	-----	2034
Db		
2161	CTGTTCAVATATGTTGGTCTCAGGTGAGTTCGTTGAATAAATCGGTTTAAAGGAGTCAAG	2220
Qy		
2035	-----GTCTCTCAGTGCAGTTGGTGAATAATCGGTTTAAAGGAGTCAAG	2078
Db		
2221	TATTTCCGCTGAATACCTTAGCTCTCTTAGTTTATGTGGTTGTAGTGATAGACAACAGG	2280
Qy		
2079	TATTTCCGCTTGAATACCTTAGCTCTCTTAGTTTATGTGGTTGTAGTGATAGACAACAGG	2138
Db		
2281	GGATCTGTACACGAGGCTTAAATTTGAAGGCGCTTTAAATATATAAATCGGTCAATA	2340
Qy		
2139	GGATCTGTACACGAGGCTTAAATTTGAAGGCGCTTTAAATATATAAATCGGTCAATA	2198
Db		
2341	GAATTTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTCAITGAC	2400
Qy		
2199	GAATTTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTCAITGAC	2258
Db		
2401	TTAGATCTGTGTGGGCATCCAGGCTGTCTATGAGAGGATACCTCTCCCTGATGGCAATTA	2460
Qy		
2259	TTAGATCTGTGTGGGCATCCAGGCTGTCTATGAGAGGATACCTCTCCCTGATGGCAATTA	2318
Db		
2461	ATGCAGAGGTGAGATATCTTTCAGGGTTGCTATTGCTGGGGCCCAAGTCACTCTGTGGATC	2520
Qy		
2319	ATGCAGAGGTGAGATATCTTTCAGGGTTGCTATTGCTGGGGCCCAAGTCACTCTGTGGATC	2378
Db		
2521	TTCTATGATACAGGATACACGGAAGCTTATATGGGTACCTCGACAGAAATGAACAGGGC	2580
Qy		
2379	TTCTATGATACAGGATACACGGAAGCTTATATGGGTACCTCGACAGAAATGAACAGGGC	2438
Db		
2581	TATTTACTTAGGATCTGTGGCCATCAACGACGAAAGTTCCCTCTGTAACCAAACTGGTTTA	2640
Qy		
2439	TATTTACTTAGGATCTGTGGCCATCAACGACGAAAGTTCCCTCTGTAACCAAACTGGTTTA	2498
Db		
2641	CTGCTCTTACATGGTTTCTCGATGAGAAATGTCCATTTTGCACATACCAAGTATATTACTG	2700
Qy		
2499	CTGCTCTTACATGGTTTCTCGATGAGAAATGTCCATTTTGCACATACCAAGTATATTACTG	2558
Db		
2701	AGTTTTTTTAGTCAGGGCTGGAAAGCCATATGATTTTACAGATCTATCCTCAGGAGAGACAC	2760
Qy		
2559	AGTTTTTTTAGTCAGGGCTGGAAAGCCATATGATTTTACAGATCTATCCTCAGGAGAGACAC	2618
Db		
2761	AGCATAAGAGTTTCTCGAATCGGGAGAACATTTATGAACCTGCATCTTTTGCACCTTACAA	2820
Qy		
2619	AGCATAAGAGTTTCTCGAATCGGGAGAACATTTATGAACCTGCATCTTTTGCACCTTACAA	2678
Db		
2821	GAAAACCTTTGGATACAGTATTTGCTCTTAAAGTGATTAATTTTTTGACCTGTGTAGAAC	2880
Qy		
2679	GAAAACCTTTGGATACAGTATTTGCTCTTAAAGTGATTAATTTTTTGACCTGTGTAGAAC	2738
Db		

RESULT 8

US-10-982-512-22

US-10-982-312-22
: Sequence 22. Application US/10982512

; sequence 22, Application OS/IOS
: publication No. US20050059081A1

: PUBLIC INFORMATION NO: 05200
: GENERAL INFORMATION:
: GENERAL INFORMATION:

GENERAL INFORMATION:
APPLICANT: Oi, Steve

APPLICANT: Q1, Steve
APPLICANT: Akinsanya, Karen

APPLICANT: Riviere, Pierre

APPLICANT: JUNIEN, JEAN-LOUIS

;
; AFFRICANI: CUNITEN, SCAR-LOU
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
;

; FILE REFERENCE: 70669

; FILE REFERENCE: 9002
; CURRENT APPLICATION NUMBER: US/10/982,512

CURRENT FILING DATE: 2004-11-05

; PRIOR APPLICATION NUMBER: US/09/976,674

; PRIOR FILING DATE: 2001-10-12

;
; PRIOR APPLICATION NUMBER: US 60/240,117

;
; PRIOR FILING DATE: 2000-10-12

; NUMBER OF SEQ ID NOS: 61

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 22

; LENGTH: 4685

TYPE: DNA

ORGANISM: *Homo sapiens*

US-10-982-512-22

Query Match 90.5%; Score 2824.4; DB 9; Length 4685;
Best Local Similarity 95.4%; Pred. No. 0;
Matches 2977; Conservative 0; Mismatches 1; Indels 142; Gaps 1;

Db 361 GCCGATACCAGAAATATCATGCTACATGATGGCTAAGGCAACCACATGATTTTCATGTTT 420
Qy 421 GTGAGAGGAATGATCCAGATGGACCTCATTTACAGACAGAAATCTATTACCTTCGCCATGCT 480
Db 421 GTGAAGAGGAATGATCCAGATGGACCTCATTTACAGACAGAAATCTATTACCTTCGCCATGCT 480
Qy 481 GGTGAGAACAGAGAAATATACACTGTTTTTATCTCGAAATTCGCCAAATCTATCAATAGAGA 540
Db 481 GGTGAGAACAGAGAAATATACACTGTTTTTATCTCGAAATTCGCCAAATCTATCAATAGAGA 540
Qy 541 GCAGTCTTAATGCTCTCTGTGAAGCCCTCTTTTGGATGCTTTTTCAGGCAACACTGGAATAT 600
Db 541 GCAGTCTTAATGCTCTCTGTGAAGCCCTCTTTTGGATGCTTTTTCAGGCAACACTGGAATAT 600
Qy 601 GGAATGTATTCGAGAGAGAAAGAACTATTAAAGAGAAAGAAACGCATTCGAAACAGTCGGA 660
Db 601 GGAATGTATTCGAGAGAGAAAGAACTATTAAAGAGAAAGAAACGCATTCGAAACAGTCGGA 660
Qy 661 ATTGCTCTTACGATTTATCAACAAAGGAAGTGGAACTATTTCTGTTTCAAGCCGCTAGTGA 720
Db 661 ATTGCTCTTACGATTTATCAACAAAGGAAGTGGAACTATTTCTGTTTCAAGCCGCTAGTGA 720
Qy 721 ATTTATCACGTAAAGATGGAGGCCCAAGAGATTTACGCAACAACTTTTAAAGGCCCAAT 780
Db 721 ATTTATCACGTAAAGATGGAGGCCCAAGAGATTTACGCAACAACTTTTAAAGGCCCAAT 780
Qy 781 CTAGTGGAAATCTAGTTGCTCCCAACATACGATGATCCCAAAATTTATGCCCGCTGATCCA 840
Db 781 CTAGTGGAAATCTAGTTGCTCCCAACATACGATGATCCCAAAATTTATGCCCGCTGATCCA 840
Qy 841 GACTCGATGCTTTTATACATAGCAACGATATTTGGATATCTAACATCTGTAACACGAGAA 900
Db 841 GACTCGATGCTTTTATACATAGCAACGATATTTGGATATCTAACATCTGTAACACGAGAA 900
Qy 901 GAAAAGAGACTCATTTATGTGCAACATGAGTAGCCCAACATGGAAGAGATCCAGATCA 960
Db 901 GAAAAGAGACTCATTTATGTGCAACATGAGTAGCCCAACATGGAAGAGATCCAGATCA 960
Qy 961 GCTGAGTGCCTACCTTTGTTCTCCAAAGAAATTTGATAGATATTTCTGGCTATTGGTGG 1020
Db 961 GCTGAGTGCCTACCTTTGTTCTCCAAAGAAATTTGATAGATATTTCTGGCTATTGGTGG 1020
Qy 1021 TGTCCAAAGCTGAAACAACTCCAGTGGTGGTAAATTTCTTAGAAATCTATATGAAGAA 1080
Db 1021 TGTCCAAAGCTGAAACAACTCCAGTGGTGGTAAATTTCTTAGAAATCTATATGAAGAA 1080
Qy 1081 AATGATGAATCTGAGGTGAAATTAATTCATGTTACATCCCTATGTTGGAACAAAGGAG 1140
Db 1081 AATGATGAATCTGAGGTGAAATTAATTCATGTTACATCCCTATGTTGGAACAAAGGAG 1140
Qy 1141 GCAGATTCATTCGTTATCCTTAAACAGGTACAGCAATCCTTAAAGTCACTTTTAAAGATG 1200
Db 1141 GCAGATTCATTCGTTATCCTTAAACAGGTACAGCAATCCTTAAAGTCACTTTTAAAGATG 1200
Qy 1201 TCAGAAATAATGATGCTGGAAGGAGGATCATAGATGCTCATAGATAAGGAATTAAT 1260
Db 1201 TCAGAAATAATGATGCTGGAAGGAGGATCATAGATGCTCATAGATAAGGAATTAAT 1260
Qy 1261 CAACCTTTGAGATCTTATTTGAAGGAGTTGAATATATTTCCAGAGCTGATGGAGCTCT 1320
Db 1261 CAACCTTTGAGATCTTATTTGAAGGAGTTGAATATATTTCCAGAGCTGATGGAGCTCT 1320
Qy 1321 GAGGAAATAATGCTTGTGCTCATCTACTAGATCGCTCCAGACTCGCCTACAGATAGTG 1380
Db 1321 GAGGAAATAATGCTTGTGCTCATCTACTAGATCGCTCCAGACTCGCCTACAGATAGTG 1380
Qy 1381 TTGATCTCACTCGAAATTTATTTCCAGTAGAAGATGATGTTATGGAAGGCGAGAGACTC 1440
Db 1381 TTGATCTCACTCGAAATTTATTTCCAGTAGAAGATGATGTTATGGAAGGCGAGAGACTC 1440
Qy 1441 ATTGAGTCAGTGCCTGATTTCTGTGACGCCCACTAATTTATCTATGAAGAAACACAGACATC 1500
Db 1441 ATTGAGTCAGTGCCTGATTTCTGTGACGCCCACTAATTTATCTATGAAGAAACACAGACATC 1500

Qy 1501 TGGATAAATATCCATGACATCTTTCTATGTTTTTCCCAAAAGTCACGAAGAGGAAATTTGAG 1560
Db 1501 TGGATAAATATCCATGACATCTTTCTATGTTTTTCCCAAAAGTCACGAAGAGGAAATTTGAG 1560
Qy 1561 TTTATTTTTGCTCTGTAATGCAAAACAGGTTTTCCGTCATTTATACAAAAATTTACATCTATT 1620
Db 1561 TTTATTTTTGCTCTGTAATGCAAAACAGGTTTTCCGTCATTTATACAAAAATTTACATCTATT 1620
Qy 1621 TTAAGGAAGCAATATATAACGATCCAGTGGTGGCTGCTGCTCCAGTGATTTCAAG 1680
Db 1621 TTAAGGAAGCAATATATAACGATCCAGTGGTGGCTGCTGCTCCAGTGATTTCAAG 1680
Qy 1681 TGTCTCTATCAAAAGAGAGATAGCAATTTACAGTGGTGAATGGGAAGTTCTTGGCCGCGAT 1740
Db 1681 TGTCTCTATCAAAAGAGAGATAGCAATTTACAGTGGTGAATGGGAAGTTCTTGGCCGCGAT 1740
Qy 1741 GGATCTTAATATCCAAAGTTGATGAAGTCAAGAGGCTGGTATATTTTGAAGGCAACAAAGAC 1800
Db 1741 GGATCTTAATATCCAAAGTTGATGAAGTCAAGAGGCTGGTATATTTTGAAGGCAACAAAGAC 1800
Qy 1801 TCCCTCTTAGAGCATCACCTGTACGTAGTACGTAAATCCTGGAGAGGTGACAAAG 1860
Db 1801 TCCCTCTTAGAGCATCACCTGTACGTAGTACGTAAATCCTGGAGAGGTGACAAAG 1860
Qy 1861 CTGACTGACCGTGGCTACTCACATTTCTGCTGCACTAGTCAGCACTGTGACTTCTTTATA 1920
Db 1861 CTGACTGACCGTGGCTACTCACATTTCTGCTGCACTAGTCAGCACTGTGACTTCTTTATA 1920
Qy 1921 AGTAAGTATAGTAACCAAGAAATCCACACTGTGTCCTTTTACAAAGCTATCAAGTCTCT 1980
Db 1921 AGTAAGTATAGTAACCAAGAAATCCACACTGTGTCCTTTTACAAAGCTATCAAGTCTCT 1980
Qy 1981 GAGATGACCCAACTTGCACAAACAAAGAAATTTTGGGCCCACTTTTGGATTCAGCAGGT 2040
Db 1981 GAGATGACCCAACTTGCACAAACAAAGAAATTTTGGGCCCACTTTTGGATTCAGCAGGT 2040
Qy 2041 CCTCTCTGACTACTACTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGAGTTTACA 2100
Db 2041 CCTCTCTGACTACTACTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGAGTTTACA 2100
Qy 2035 ----- 2034
Db 2035 ----- 2034
Qy 2101 TTGTATGGGATGCTCTACAGGCTCTGATCTACAGGCTGGAAAGAAATATCTCTACTGTG 2160
Db 2035 ----- 2034
Qy 2161 CTGTTTCATATATGGTGGTCTCTCAGGTGCAATGATGATTAATCGGTTTAAAGGAGTCAAG 2220
Db 2035 ----- 2078
Qy 2221 TATTTCCGCTTGAATACCTTAGCTCTCTAGGTTATGTTGGTTGATGATAGACAAAG 2280
Db 2079 TATTTCCGCTTGAATACCTTAGCTCTCTAGGTTATGTTGGTTGATGATAGACAAAG 2138
Qy 2281 GGATCCTGTCAACGAGGCTTAAATTTGAAGGCGCTTTAAATATAAAATGGGTCAAATA 2340
Db 2139 GGATCCTGTCAACGAGGCTTAAATTTGAAGGCGCTTTAAATATAAAATGGGTCAAATA 2198
Qy 2341 GAAATTTGACGATCAGGTGGAAGGACTCCAAATATCTAGCTCTCGATATGATTTCAATGAC 2400
Db 2199 GAAATTTGACGATCAGGTGGAAGGACTCCAAATATCTAGCTCTCGATATGATTTCAATGAC 2258
Qy 2401 TTAGATCTGCTGGGATCCAGGCTGCTCTATGGAGGATACCTCTCCCTGATGGGATTA 2460
Db 2259 TTAGATCTGCTGGGATCCAGGCTGCTCTATGGAGGATACCTCTCCCTGATGGGATTA 2318
Qy 2461 ATGACAGGTGATGATCTTCAAGGTTGCTATTCGCTGGGCGCCAGTCACTCTGTGGATC 2520
Db 2319 ATGACAGGTGATGATCTTCAAGGTTGCTATTCGCTGGGCGCCAGTCACTCTGTGGATC 2378
Qy 2521 TTCTATGATACAGGATACACGGAACGTTATATGGGTCACTCTGACCAAGATGAACAGGCG 2580
Db 2379 TTCTATGATACAGGATACACGGAACGTTATATGGGTCACTCTGACCAAGATGAACAGGCG 2438

```
Qy 2581 TATTACTTAGGATCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGGAACCAATCGTTTA 2640
Db |||||
Qy 2439 TATTACTTAGGATCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGGAACCAATCGTTTA 2498
Db |||||
Qy 2641 CTGCTCTTACATGGTTTCTCGATGAGATGTCCATTTTGGACATACAGTATTAATCTG 2700
Db |||||
Qy 2499 CTGCTCTTACATGGTTTCTCGATGAGATGTCCATTTTGGACATACAGTATTAATCTG 2558
Db |||||
Qy 2701 AGTTTTTTAGTAGGGCTCGAAAGCCATATGATTTACAGATCTATCTCTCAGGAGAGACAC 2760
Db |||||
Qy 2559 AGTTTTTTAGTAGGGCTCGAAAGCCATATGATTTACAGATCTATCTCTCAGGAGAGACAC 2618
Db |||||
Qy 2761 AGCATAGAGTTTCTCGAATCGGAGAGACATTATGAACTGCATCTTTTGGACTACCTTCAA 2820
Db |||||
Qy 2619 AGCATAGAGTTTCTCGAATCGGAGAGACATTATGAACTGCATCTTTTGGACTACCTTCAA 2678
Db |||||
Qy 2821 GAAAACTTGGATCAGTATTTGCTGTCTTAAAGTGATATAATTTTGGACTGTGTAGAAC 2880
Db |||||
Qy 2679 GAAAACTTGGATCAGTATTTGCTGTCTTAAAGTGATATAATTTTGGACTGTGTAGAAC 2738
Db |||||
Qy 2881 TCTCTGTATACACTGGCTATTTAAACCAATGAGGAGTTTAAATCAACAGAAACACAGA 2940
Db |||||
Qy 2739 TCTCTGTATACACTGGCTATTTAAACCAATGAGGAGTTTAAATCAACAGAAACACAGA 2798
Db |||||
Qy 2941 ATTGATCATCACATTTTGTATCTGCCATGTAAACATCTACTCTCGAATAAATTTGGTG 3000
Db |||||
Qy 2799 ATTGATCATCACATTTTGTATCTGCCATGTAAACATCTACTCTCGAATAAATTTGGTG 2858
Db |||||
Qy 3001 CCATGACAGGGTCTACGGTTTGTGTAGTAATCTAATACCTTAACCCCAATGCTCAAAA 3060
Db |||||
Qy 2859 CCATGACAGGGTCTACGGTTTGTGTAGTAATCTAATACCTTAACCCCAATGCTCAAAA 2918
Db |||||
Qy 3061 TCAATGATACATATTCCTGAGAGACCCAGCAATACCAATAGATTACTTAAAAA 3120
Db |||||
Qy 2919 TCAATGATACATATTCCTGAGAGACCCAGCAATACCAATAGATTACTTAAAAA 2978
Db |||||
```

RESULT 9

```
US-09-976-674-20
; Sequence 20, Application US/09976674
; Patent No. US20020115843A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
; LENGTH: 4676
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-976-674-20
```

```
Query Match 89.9%; Score 2806.4; DB 3; Length 4676;
Best Local Similarity 95.1%; Pred. No. 0;
Matches 2968; Conservative 0; Mismatches 1; Indels 151; Gaps 1;

Qy 1 AAGTGCTAAAGCCTCCGAGGCCAAGGCCGTGCTACTGCGCGCGTCTTCTTAGTGCCG 60
Db |||||
Qy 1 AAGTGCTAAAGCCTCCGAGGCCAAGGCCGTGCTACTGCGCGCGTCTTCTTAGTGCCG 60
Db |||||
Qy 61 CGTTGCGCGCGCTGGTTGTTCACCGCGCGCGCGCGAGGAGCCACTGCAACACGAGGACCG 120
Db |||||
Qy 61 CGTTGCGCGCGCTGGTTGTTCACCGCGCGCGCGCGAGGAGCCACTGCAACACGAGGACCG 120
Db |||||
Qy 121 GAGTGGAGGGCGGCGACGATGAAGGGGGCGGCGCGCTCCATAGCGCAGCTCGGGACGG 180
Db |||||
```

```
Db 121 GAGTGGAGGGCGGCGACGATGAAGGGGGCGGCGCGCTCCATAGCGCAGCTCGGGACGG 180
Qy |||||
Db 181 TCCGGGCGGGCGGCGGAGGAAATGCAACATGCGCAGCAGCAATGGAACAGAAACAG 240
Qy |||||
Db 181 TCCGGGCGGGCGGCGGAGGAAATGCAACATGCGCAGCAGCAATGGAACAGAAACAG 240
Qy |||||
Qy 241 CTGGGTGTTGAGATATTTGAAACTCGGCACTGTGAGGAGATATTTGAATCACAGGATCGG 300
Db |||||
Qy 241 CTGGGTGTTGAGATATTTGAAACTCGGCACTGTGAGGAGATATTTGAATCACAGGATCGG 300
Db |||||
Qy 301 CCTAAATCTGAGCGCTTTTATGTTGAGCGGTATTCCTGAGTCAAGCTTAAAGCTGCTT 360
Db |||||
Qy 301 CCTAAATCTGAGCGCTTTTATGTTGAGCGGTATTCCTGAGTCAAGCTTAAAGCTGCTT 360
Db |||||
Qy 361 GCCGATACACAGAAATATCATGGCTACATGCTAAGGCAACCAATGATTTTCAATGTTT 420
Db |||||
Qy 361 GCCGATACACAGAAATATCATGGCTACATGCTAAGGCAACCAATGATTTTCAATGTTT 420
Db |||||
Qy 421 GTGAAGAGGAATGATCCAGATGGACCTCATTTAGAGAGATCTATTTACCTTGGCTGCT 480
Db |||||
Qy 421 GTGAAGAGGAATGATCCAGATGGACCTCATTTAGAGAGATCTATTTACCTTGGCTGCT 480
Db |||||
Qy 481 GGTGAGAACAGAGAAATACACTGTTTATCTGAAATTTCCAAACCTATCAATAGAGCA 540
Db |||||
Qy 481 GGTGAGAACAGAGAAATACACTGTTTATCTGAAATTTCCAAACCTATCAATAGAGCA 540
Db |||||
Qy 541 GCAGTCTTAATGCTCTCTTGGAAAGCTCTTTTGGATCTTTTTCAGGCAACACTGAGCTAT 600
Db |||||
Qy 541 GCAGTCTTAATGCTCTCTTGGAAAGCTCTTTTGGATCTTTTTCAGGCAACACTGAGCTAT 600
Db |||||
Qy 601 GGAATGTATTTCTCGAGAGAGAACTATTAAGAGAGAAAGAAACGCAATTTGGAACAGTCGGA 660
Db |||||
Qy 601 GGAATGTATTTCTCGAGAGAGAACTATTAAGAGAGAAAGAAACGCAATTTGGAACAGTCGGA 660
Db |||||
Qy 661 ATTGCTCTTTACGATTTATCACCAAGAGTGGAACTATTTCTGTTTCAAGCCGCTAGTGA 720
Db |||||
Qy 661 ATTGCTCTTTACGATTTATCACCAAGAGTGGAACTATTTCTGTTTCAAGCCGCTAGTGA 720
Db |||||
Qy 721 ATTTATCACGCTAAAGATGAGGGCCACAAAGATTTTACGCAACCACTTTTAAAGCCCAAT 780
Db |||||
Qy 721 ATTTATCACGCTAAAGATGAGGGCCACAAAGATTTTACGCAACCACTTTTAAAGCCCAAT 780
Db |||||
Qy 781 CTAGTGGAAACTAGTTGTCCCAACATACGATGGATCCAAATTTATGCCCCGCTGATCCA 840
Db |||||
Qy 781 CTAGTGGAAACTAGTTGTCCCAACATACGATGGATCCAAATTTATGCCCCGCTGATCCA 840
Db |||||
Qy 841 GACTGGATTCCTTTATACATAGCAACGATATTTGGATATCTTAACATCTGAACACAGAA 900
Db |||||
Qy 841 GACTGGATTCCTTTATACATAGCAACGATATTTGGATATCTTAACATCTGAACACAGAA 900
Db |||||
Qy 901 GAAAGGAGACTCACTTATGTCACATGAGCTAGCCCAATGGAAGAGATGCCAGATCA 960
Db |||||
Qy 901 GAAAGGAGACTCACTTATGTCACATGAGCTAGCCCAATGGAAGAGATGCCAGATCA 960
Db |||||
Qy 961 GCTGAGTCTGCTACCTTTGTTCTCCAAAGAGAAATTTGATAGATAATCTGGCTATTTGGTGG 1020
Db |||||
Qy 961 GCTGAGTCTGCTACCTTTGTTCTCCAAAGAGAAATTTGATAGATAATCTGGCTATTTGGTGG 1020
Db |||||
Qy 1021 TGTCCTAAAGCTGAAACCACTCCAGTGTGGTGAATTTCTTAGAATTTCTATATGAAGAA 1080
Db |||||
Qy 1021 TGTCCTAAAGCTGAAACCACTCCAGTGTGGTGAATTTCTTAGAATTTCTATATGAAGAA 1080
Db |||||
Qy 1081 AATGATGAATCTGAGTGGGAAATTTATGTTTATCATCTCCCTATCTGCGAAACAGGAGG 1140
Db |||||
Qy 1081 AATGATGAATCTGAGTGGGAAATTTATGTTTATCATCTCCCTATCTGTTGAAACAGGAGG 1140
Db |||||
Qy 1141 GCAGATTCATTTCCGTTATCTTAAACAGGTCAGCAAAATCTTAAAGTCACTTTTAAAGTG 1200
Db |||||
Qy 1141 GCAGATTCATTTCCGTTATCTTAAACAGGTCAGCAAAATCTTAAAGTCACTTTTAAAGTG 1200
Db |||||
Qy 1201 TCAGAAATATGATTTGATGCTGAAGGAGGATCATAGATGTCATAGTAAGGAACTAAT 1260
Db |||||
```


Db	1201	TCAGAAATATGATGTGATGCTGAAGGAGGATCATATAGATGTCTATAGATTAAGAACTAATT	1260
Qy	1261	CAACCTTTTGAGATCTTATTTGAAGGAGTTGAATATATTTGCGAGAGCTGGATGAGCTCCT	1320
Db	1261	CAACCTTTTGAGATCTTATTTGAAGGAGTTGAATATATTTGCGAGAGCTGGATGAGCTCCT	1320
Qy	1321	GAGGAAATATGCTTGGTCCATCCTACTAGATCGCTCCAGACTCGCTCAAGATAGTG	1380
Db	1321	GAGGAAATATGCTTGGTCCATCCTACTAGATCGCTCCAGACTCGCTCAAGATAGTG	1380
Qy	1381	TTGATCTCACCTGAAATATTTATCCAGTAGAAGATGATGTTATGGAAAGGCAGAGACTC	1440
Db	1381	TTGATCTCACCTGAAATATTTATCCAGTAGAAGATGATGTTATGGAAAGGCAGAGACTC	1440
Qy	1441	ATTGAGTCAGTGCCTGATCTGTGAGCCCACTAAATATCTATGAAGAAACACAGACATC	1500
Db	1441	ATTGAGTCAGTGCCTGATCTGTGAGCCCACTAAATATCTATGAAGAAACACAGACATC	1500
Qy	1501	TGGATAAATATCCATGACATCTTTCAATGTTTTTCCCAAGTCACGAGAGGAAATTCAG	1560
Db	1501	TGGATAAATATCCATGACATCTTTCAATGTTTTTCCCAAGTCACGAGAGGAAATTCAG	1560
Qy	1561	TTTATTTTGGCTCTGAAATGCAAAACAGGTTTCGGTCATTTATACAAATTTACATCTATT	1620
Db	1561	TTTATTTTGGCTCTGAAATGCAAAACAGGTTTCGGTCATTTATACAAATTTACATCTATT	1620
Qy	1621	TTAAAGGAAAGCAATATAAAGATTAACAGATCCAGTGGTGGCTGCTCTCCAGTGATTTCAAG	1680
Db	1621	TTAAAGGAAAGCAATATAAAGATTAACAGATCCAGTGGTGGCTGCTCTCCAGTGATTTCAAG	1680
Qy	1681	TGCTCTATCAAGAGGAGATAGCAATTAACAGTGGTGAATGGGAAGTTCTTGGCGGGAT	1740
Db	1681	TGCTCTATCAAGAGGAGATAGCAATTAACAGTGGTGAATGGGAAGTTCTTGGCGGGAT	1740
Qy	1741	GGATCTAAATATCCAAAGTTGATGAAGTCAGAAAGTCGGTATATTTTGAAGGCCACAAAGAC	1800
Db	1741	GGATCTAAATATCCAAAGTTGATGAAGTCAGAAAGTCGGTATATTTTGAAGGCCACAAAGAC	1800
Qy	1801	TCCCTTTTAGACATCACTCTGATAGTACATTTCTGTGATCAGTTACGTAAATCCTCGAGAGGTGCAAGG	1860
Db	1801	TCCCTTTTAGACATCACTCTGATAGTACATTTCTGTGATCAGTTACGTAAATCCTCGAGAGGTGCAAGG	1860
Qy	1861	CTGACTGACCGTGGCTACTCAATTTCTGTGATCAGTCAGTCAGTCACTGTGACTTCTTTATA	1920
Db	1861	CTGACTGACCGTGGCTACTCAATTTCTGTGATCAGTCAGTCAGTCACTGTGACTTCTTTATA	1920
Qy	1921	AGTAAGTATAGTAACCCAGAGATCCACACTGTGTGTCCTTTTACAAGCTATCAAGTCTCT	1980
Db	1921	AGTAAGTATAGTAACCCAGAGATCCACACTGTGTGTCCTTTTACAAGCTATCAAGTCTCT	1980
Qy	1981	GAAGATGCCCAACTTGCMAACAAAGGAATTTTGGGCCACCAATTTTGGATTCAGCAGGT	2040
Db	1981	GAAGATGCCCAACTTGCMAACAAAGGAATTTTGGGCCACCAATTTTGGATTCAGCAGGT	2040
Qy	2041	CCTCTTCTGACTATCTCTCCAGAAATTTTCTTTTGAAGTACTACTGGATTTACA	2100
Db	2041	CCTCTTCTGACTATCTCTCCAGAAATTTTCTTTTGAAGTACTACTGGATTTACA	2100
Qy	2101	TTGTATGGATGCTCTACAAGCCTCATGATCTACAGCTGGAAGAAATATCTACTGTG	2160
Db	2101	TTGTATGGATGCTCTACAAGCCTCATGATCTACAGCTGGAAGAAATATCTACTGTG	2160
Qy	2161	CTGTTTCATATATGTTGGTCTCTCAGGTGCGATTTGGTGAATATCGTTTAAAGGATCAAG	2220
Db	2161	CTGTTTCATATATGTTGGTCTCTCAGGTGCGATTTGGTGAATATCGTTTAAAGGATCAAG	2220
Qy	2221	TATTTCCGCTTGAATACCCCTAGCCCTCTCTAGGTTATGTGTTGTAGTATAGACACAGG	2280
Db	2281	GGATCCTGTACCGGGGCTTAAATTTGAAGCGGCTTTAAATATAAAATGGGTCAAATA	2340
Db	2180	-----GGGTCAAATA	2180

Qy	2341	GAAATTCAGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTCAATTGAC	2400
Db	2190	GAAATTCAGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTCAATTGAC	2249
Qy	2401	TTAGATCGTGGGCGATCCACGGCTGGTCTCTATGGAGGATACCTCTCCCTGATGCAATTA	2460
Db	2250	TTAGATCGTGGGCGATCCACGGCTGGTCTCTATGGAGGATACCTCTCCCTGATGCAATTA	2309
Qy	2461	ATGCAAGGTCAGATATCTTTCAGGGTTGCTATTTGCTGGGGCCCGAGTCACTCTGGATC	2520
Db	2310	ATGCAAGGTCAGATATCTTTCAGGGTTGCTATTTGCTGGGGCCCGAGTCACTCTGGATC	2369
Qy	2521	TTCTATGATACAGGATACACGGAAACGTTATATGGGTTCACCCCTGACAGAAATGAAACGGGC	2580
Db	2370	TTCTATGATACAGGATACACGGAAACGTTATATGGGTTCACCCCTGACAGAAATGAAACGGGC	2429
Qy	2581	TATTACTTTAGGATCTGTGGCCATGCAAGCAGAAAAAGTTCCCTCTGAAACCAATCGTTTA	2640
Db	2430	TATTACTTTAGGATCTGTGGCCATGCAAGCAGAAAAAGTTCCCTCTGAAACCAATCGTTTA	2489
Qy	2641	CTGCTCTTACATGGTTTCTCTGGATGGAATGTCCATTTTGCACATACCAAGTATATTACTG	2700
Db	2490	CTGCTCTTACATGGTTTCTCTGGATGGAATGTCCATTTTGCACATACCAAGTATATTACTG	2549
Qy	2701	AGTTTTTTAGTGAGGGCTGGAAGCCATATGATTTTACAGATCTATCTCTCAGGAGAGACAC	2760
Db	2550	AGTTTTTTAGTGAGGGCTGGAAGCCATATGATTTTACAGATCTATCTCTCAGGAGAGACAC	2609
Qy	2761	AGCATAAGAGTTCTCTGAATCGGAGAACATTTATGAACTGTCATCTTTTGCACTACCTTCAA	2820
Db	2610	AGCATAAGAGTTCTCTGAATCGGAGAACATTTATGAACTGTCATCTTTTGCACTACCTTCAA	2669
Qy	2821	GAAACCTTGGATCAGCTATTTGCTCTAAAGTGTATATAATTTTGAAGTGTATAGAAC	2880
Db	2670	GAAACCTTGGATCAGCTATTTGCTCTAAAGTGTATATAATTTTGAAGTGTATAGAAC	2729
Qy	2881	TCTCTGGTATACATCGGCTATTTAAACCAATGAGGAGTTTAAATCAACAGAAACACAGA	2940
Db	2730	TCTCTGGTATACATCGGCTATTTAAACCAATGAGGAGTTTAAATCAACAGAAACACAGA	2789
Qy	2941	ATGATCATCATATTTGATACCTGCGCATGTAACTACTCTCTGAAATTAATGTGGTG	3000
Db	2790	ATGATCATCATATTTGATACCTGCGCATGTAACTACTCTCTGAAATTAATGTGGTG	2849
Qy	3001	CCATGAGGGGTCTACGGTTTGTGGTAGTAATCTTAATACCTTAAACCCCATGCTCAAAA	3060
Db	2850	CCATGAGGGGTCTACGGTTTGTGGTAGTAATCTTAATACCTTAAACCCCATGCTCAAAA	2909
Qy	3061	TCAATGATACATATTCCTGAGAGACCCAGCAATACCAATTAAGAAATTAATAAAAAAAA	3120
Db	2910	TCAATGATACATATTCCTGAGAGACCCAGCAATACCAATTAAGAAATTAATAAAAAAAA	2969

RESULT 10

US-10-982-512-20

Sequence 20, Application US/10982512

Publication No. US20050059081A1

GENERAL INFORMATION:

APPLICANT: Qi, Steve

APPLICANT: Akinsanya, Karen

APPLICANT: Riviere, Pierre

APPLICANT: Jannin, Jean-Louis

TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV

FILE REFERENCE: 70669

CURRENT APPLICATION NUMBER: US/10/982, 512

CURRENT FILING DATE: 2004-11-05

PRIOR APPLICATION NUMBER: US/09/976, 674

PRIOR FILING DATE: 2001-10-12

PRIOR APPLICATION NUMBER: US 60/240,117

PRIOR FILING DATE: 2000-10-12

NUMBER OF SEQ ID NOS: 61

SOFTWARE: PatentIn version 3.1

! SEQ ID NO 20
! LENGTH: 4676
! TYPE: DNA
! ORGANISM: Homo sapiens
US-10-982-512-20

Query Match 89.9%; Score 2806.4; DB 9; Length 4676;
Best Local Similarity 95.1%; Pred. No. 0;
Matches 2968; Conservative 0; Mismatches 1; Indels 151; Gaps 1;

Qy	1	AAGTGTCTAAGAGCTCCGAGGCCAAGCGCTGCTACTGCGCGCTGCTCTTCTTAGTGCCG	60
Db	1	AAGTGTCTAAGAGCTCCGAGGCCAAGCGCTGCTACTGCGCGCTGCTCTTCTTAGTGCCG	60
Qy	61	CGTTCCGCCCTGGGTTGTCAACCGCGCGCGCGCGAGGAGCCACATGCAACAGGACCG	120
Db	61	CGTTCCGCCCTGGGTTGTCAACCGCGCGCGCGCGAGGAGCCACATGCAACAGGACCG	120
Qy	121	GAGTGAGGCGGCGAGCATGAAAGCGGCGCGCGCGCTCCATAGCGCATGTCGGGACCG	180
Db	121	GAGTGAGGCGGCGAGCATGAAAGCGGCGCGCGCGCTCCATAGCGCATGTCGGGACCG	180
Qy	181	TCCGCGCGGCGCGGAGGAAATGCAACATGGCAGCAGCAATGGAAACAGAACAG	240
Db	181	TCCGCGCGGCGCGGAGGAAATGCAACATGGCAGCAGCAATGGAAACAGAACAG	240
Qy	241	CTGGGTGTTGAGATATTTGAAACTGCGACTGTGAGGAGAAATTTGAATCACAGGATCGG	300
Db	241	CTGGGTGTTGAGATATTTGAAACTGCGACTGTGAGGAGAAATTTGAATCACAGGATCGG	300
Qy	301	CCTAAATTTGGAGCTTTTATGTTGAGCGGTATTCCTGGAGTCAAGTTAAAAAGCTGCTT	360
Db	301	CCTAAATTTGGAGCTTTTATGTTGAGCGGTATTCCTGGAGTCAAGTTAAAAAGCTGCTT	360
Qy	361	GCCGATACAGAAATATCATGCTACATGATGGCTAGGACACACATGATTTTCATGTTT	420
Db	361	GCCGATACAGAAATATCATGCTACATGATGGCTAGGACACACATGATTTTCATGTTT	420
Qy	421	GTCAAGAGGAATGATCCAGATGGAATCTCATTGAGAGAAATCTATTACCTTGGCATGCT	480
Db	421	GTCAAGAGGAATGATCCAGATGGAATCTCATTGAGAGAAATCTATTACCTTGGCATGCT	480
Qy	481	GGTGAGAACAGAGAAATACACTGTTTATCTGAAATTTCCAAAACTCAATAGAGCA	540
Db	481	GGTGAGAACAGAGAAATACACTGTTTATCTGAAATTTCCAAAACTCAATAGAGCA	540
Qy	541	GCAGTCTTAATGCTCTCTTGGAGAGCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT	600
Db	541	GCAGTCTTAATGCTCTCTTGGAGAGCTCTTTTGGATCTTTTTCAGGCAACACTGGACTAT	600
Qy	601	GGATGTATTTCTGAGAGAGAACTATTTAAGAGAGAAAGCAACGATTTGGAACAGTCGGA	660
Db	601	GGATGTATTTCTGAGAGAGAACTATTTAAGAGAGAAAGCAACGATTTGGAACAGTCGGA	660
Qy	661	ATTGCTCTTACGATTTATCACAAAGAGTGGAAATTTCTGTTTCAAGCCGGTAGTGGGA	720
Db	661	ATTGCTCTTACGATTTATCACAAAGAGTGGAAATTTCTGTTTCAAGCCGGTAGTGGGA	720
Qy	721	ATTATATCACTAAAGATGGAGGCGCAAGGATTTACGCAACAACTTTAAGGCCCAAT	780
Db	721	ATTATATCACTAAAGATGGAGGCGCAAGGATTTACGCAACAACTTTAAGGCCCAAT	780
Qy	781	CTAGTGGAAACTAGTTGTCCCAACATACGATGGATCCAAATTTATGCCCGCTGATCCA	840
Db	781	CTAGTGGAAACTAGTTGTCCCAACATACGATGGATCCAAATTTATGCCCGCTGATCCA	840
Qy	841	GACTGGATTCCTTTTATACATAGCAACGATATTTGGATATCTAAACATCGTAAACAGAGAA	900
Db	841	GACTGGATTCCTTTTATACATAGCAACGATATTTGGATATCTAAACATCGTAAACAGAGAA	900
Qy	901	GAAAGGAGACTCATTTATGTGCAAAATGAGCTAGCCAAATGGAGAGAGATGCCAGATCA	960
Db	901	GAAAGGAGACTCATTTATGTGCAAAATGAGCTAGCCAAATGGAGAGAGATGCCAGATCA	960

Qy	961	GCTGGAGTCCGCTACCTTTGTTCTCCAGAGAGAAATTTGATAGATATTTCTGGCTATTTGGTG	1020
Db	961	GCTGGAGTCCGCTACCTTTGTTCTCCAGAGAGAAATTTGATAGATATTTCTGGCTATTTGGTG	1020
Qy	1021	TGTCCAAAAGCTGAAACAACTCCAGTGGTGTAAATTTCTTAGAATTTCTATATGAAGAA	1080
Db	1021	TGTCCAAAAGCTGAAACAACTCCAGTGGTGTAAATTTCTTAGAATTTCTATATGAAGAA	1080
Qy	1081	AATGATGAATCTGAGGTGGAAATTTATCATGTTACATCCCCTATGTTGGAACAAAGGAGG	1140
Db	1081	AATGATGAATCTGAGGTGGAAATTTATCATGTTACATCCCCTATGTTGGAACAAAGGAGG	1140
Qy	1141	GCAGATTCATTCGGTTATCTTAAACAGGTACAGCAAAATCTTAAAGTCACTTTTAAAGATG	1200
Db	1141	GCAGATTCATTCGGTTATCTTAAACAGGTACAGCAAAATCTTAAAGTCACTTTTAAAGATG	1200
Qy	1201	TCAGAAATATGATTCGATGCTGAAGGAGGATCATAGATGTCATAGATTAAGAACTAAT	1260
Db	1201	TCAGAAATATGATTCGATGCTGAAGGAGGATCATAGATGTCATAGATTAAGAACTAAT	1260
Qy	1261	CAACCTTTTGAGATTTCTATTTGAAGGAGTTGAATATTTGCCAGAGCTGGATGACTCCT	1320
Db	1261	CAACCTTTTGAGATTTCTATTTGAAGGAGTTGAATATTTGCCAGAGCTGGATGACTCCT	1320
Qy	1321	GAGGAAAAATATGCTTGGTCCATCTCTAGATTCGCTCCAGACTCGCTCAGATAGTG	1380
Db	1321	GAGGAAAAATATGCTTGGTCCATCTCTAGATTCGCTCCAGACTCGCTCAGATAGTG	1380
Qy	1381	TTGATCTCACTGAATTTATTTCCAGTAGAAGATGATTTATGGAAGGCGAGAGATC	1440
Db	1381	TTGATCTCACTGAATTTATTTCCAGTAGAAGATGATTTATGGAAGGCGAGAGATC	1440
Qy	1441	ATTGAGTCACTGCTGATTTCTGAGCGCACTAAATTTCTATGAAGAAACAAACAGATC	1500
Db	1441	ATTGAGTCACTGCTGATTTCTGAGCGCACTAAATTTCTATGAAGAAACAAACAGATC	1500
Qy	1501	TGGATAAATATTCATGACATCTTTTCATGTTTCCCAAGTCAAGAGAGAAATTTGAG	1560
Db	1501	TGGATAAATATTCATGACATCTTTTCATGTTTCCCAAGTCAAGAGAGAGAAATTTGAG	1560
Qy	1561	TTTATTTTTCCTCTGATGCAAAACAGGTTTCCGTCATTTTATACAAATTTACATCTATT	1620
Db	1561	TTTATTTTTCCTCTGATGCAAAACAGGTTTCCGTCATTTTATACAAATTTACATCTATT	1620
Qy	1621	TTAAAGGAAACAAATATATAAAGATCCAGTGGTGGCTGCTGCTGCTCCTCAAGTATTTCAAG	1680
Db	1621	TTAAAGGAAACAAATATATAAAGATCCAGTGGTGGCTGCTGCTGCTCCTCAAGTATTTCAAG	1680
Qy	1681	TGCTCTATCAAGAGGAGATAGCAATTTACAGTGGTGGAAATGGGAAGTTCTTGGCCGGCAT	1740
Db	1681	TGCTCTATCAAGAGGAGATAGCAATTTACAGTGGTGGAAATGGGAAGTTCTTGGCCGGCAT	1740
Qy	1741	GGATCTAATATCCAAAGTGGATGAAATGAGAGTGGATATTTTGAAGGCAACCAAGAC	1800
Db	1741	GGATCTAATATCCAAAGTGGATGAAATGAGAGTGGATATTTTGAAGGCAACCAAGAC	1800
Qy	1801	TCCCTTTTAGGAGCATCCCTGATAGTACGTAGTACGTAAATCCCTGGAGAGGTGACAAGG	1860
Db	1801	TCCCTTTTAGGAGCATCCCTGATAGTACGTAGTACGTAAATCCCTGGAGAGGTGACAAGG	1860
Qy	1861	CTGACTCAACCGTGGCTACTCACTTCTGCTGATCAGTCACTGTCAGTCTGTTCTTTTATA	1920
Db	1861	CTGACTCAACCGTGGCTACTCACTTCTGCTGATCAGTCACTGTCAGTCTGTTCTTTTATA	1920
Qy	1921	AGTAAGTATAGTACCAAGAGAAATCCACTGTGTGTCCTTTTACAGCTATCAAGTCTCT	1980
Db	1921	AGTAAGTATAGTACCAAGAGAAATCCACTGTGTGTCCTTTTACAGCTATCAAGTCTCT	1980
Qy	1981	GAAAGTACCCCACTTGCBAACAAAGGAATTTTGGGCCACCATTTTGGATTTACAGAGT	2040
Db	1981	GAAAGTACCCCACTTGCBAACAAAGGAATTTTGGGCCACCATTTTGGATTTACAGAGT	2040

Qy 2041 CTTCTTCTGTACTATACCTCTCCAGAAATTTCTCTTTTGAAGTACTACTGGATTAC 2100
Db 2041 CTTCTTCTGTACTATACCTCTCCAGAAATTTCTCTTTTGAAGTACTACTGGATTAC 2100
Qy 2101 TTGTATGGGATGCTCTCAAGCCTCATGATCTACAGCCTGGAAAGAAATATCTCTACTGTG 2160
Db 2101 TTGTATGGGATGCTCTCAAGCCTCATGATCTACAGCCTGGAAAGAAATATCTCTACTGTG 2160
Qy 2161 CTGTTTCATATATGGTGGTCTCAGGTGCAGTGTGGTGAATAATCGGTTTAAAGGAGTCAAG 2220
Db 2161 CTGTTTCATATATGGTGGTCTCAGGTGCAGTGTGGTGAATAATCGGTTTAAAGGAGTCAAG 2220
Qy 2221 TATTTCCGCTTGAATACCTAGCCTCTCTAGGTTATGTGGTTGTAGTGATAGACACAGG 2280
Db 2180 ----- 2179
Qy 2281 GGATCCTGTCCAGGCGCTTAAATTTGAGGCGCCTTTAAATATAAAATGGTCAATA 2340
Db 2180 -----GGTCAATA 2189
Qy 2341 GAAATGAGATCAGGTGGAGGACTCCAATATCTAGCTTCTCGATATGATTTTCATTGAC 2400
Db 2190 GAAATGAGATCAGGTGGAGGACTCCAATATCTAGCTTCTCGATATGATTTTCATTGAC 2249
Qy 2401 TTAGATCGTGTGGGATCCAGCGTGTGCTCTATGGAGATACCTCTCCCTGATGGCATTA 2460
Db 2250 TTAGATCGTGTGGGATCCAGCGTGTGCTCTATGGAGATACCTCTCCCTGATGGCATTA 2309
Qy 2461 ATGCAAGGTTCAGATATCTTCAGGGTTGCTATTTGCTGGGGCCCACTCATCTCTGTGGATC 2520
Db 2310 ATGCAAGGTTCAGATATCTTCAGGGTTGCTATTTGCTGGGGCCCACTCATCTCTGTGGATC 2369
Qy 2521 TTCTATGATACAGGATACACGGAACTGTTATATGGGTCACTCCCTGACCAAGATGAACAGGGC 2580
Db 2370 TTCTATGATACAGGATACACGGAACTGTTATATGGGTCACTCCCTGACCAAGATGAACAGGGC 2429
Qy 2581 TATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGAAACCAATCGTTTA 2640
Db 2430 TATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCCTCTGAAACCAATCGTTTA 2489
Qy 2641 CTGCTCTTACATGGTTTCTCTGATGAGATGTCATTTTGACATACCAAGTATATACATG 2700
Db 2490 CTGCTCTTACATGGTTTCTCTGATGAGATGTCATTTTGACATACCAAGTATATACATG 2549
Qy 2701 AGTATTTTGTGAGGCTGGAAGCCATATGATTTTACAGATCTATCTCAGGAGAGACAC 2760
Db 2550 AGTATTTTGTGAGGCTGGAAGCCATATGATTTTACAGATCTATCTCAGGAGAGACAC 2609
Qy 2761 AGCATAGATTTCTGAAATCGGAGAACATTAATGATGATCTTTTGGCACTACCTTCAA 2820
Db 2610 AGCATAGATTTCTGAAATCGGAGAACATTAATGATGATCTTTTGGCACTACCTTCAA 2669
Qy 2821 GAAAACTTGGATCAGTATTTGCTCTCTTAAAGTATATTTTGGACCTGTGTAGAAC 2880
Db 2670 GAAAACTTGGATCAGTATTTGCTCTCTTAAAGTATATTTTGGACCTGTGTAGAAC 2729
Qy 2881 TCTCTGGTATACATGGCTATTTTAAACCAATGAGAGGTTTAAATCAACAGAAACACAGA 2940
Db 2730 TCTCTGGTATACATGGCTATTTTAAACCAATGAGAGGTTTAAATCAACAGAAACACAGA 2789
Qy 2941 ATTGATCATCATATTTGATACCTGCCATGTAAACATCTACTCTCTGAAAAATAATGTGGTG 3000
Db 2790 ATTGATCATCATATTTGATACCTGCCATGTAAACATCTACTCTCTGAAAAATAATGTGGTG 2849
Qy 3001 CCATGAGGGGCTACGGTTTGTGTGTAGTAACTTAAATACCTTTAAACCCACATGCTCAAAA 3060
Db 2850 CCATGAGGGGCTACGGTTTGTGTGTAGTAACTTAAATACCTTTAAACCCACATGCTCAAAA 2909
Qy 3061 TCAAATGATACATATTTCTGAGAGAGCCAGCAATACCATAGAAATTAATAAAAAA 3120
Db 2910 TCAAATGATACATATTTCTGAGAGAGCCAGCAATACCATAGAAATTAATAAAAAA 2969

RESULT 11
US-10-170-789-39
; Sequence 39, Application US/10170789
; Publication No. US20030180930A1
; GENERAL INFORMATION:
; APPLICANT: Rachel E. Meyers
; APPLICANT: Olandt, Peter J.
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Williamson, Mark
; APPLICANT: Welch, Nadine
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASE, PHOSPHATASE,
; FILE REFERENCE: 10448-191001
; CURRENT APPLICATION NUMBER: US/10/170,789
; CURRENT FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: US 09/797,039
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: PCT/US01/06525
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/186,061
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 09/882,166
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: PCT/US01/19269
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/212,078
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/934,406
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: PCT/US01/26052
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/226,740
; PRIOR FILING DATE: 2000-08-21
; PRIOR APPLICATION NUMBER: US 09/861,801
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: PCT/US01/16549
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: US 60/205,508
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: PCT/US01/07138
; PRIOR FILING DATE: 2001-03-06
; PRIOR APPLICATION NUMBER: PCT/US01/07138
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: US 60/187,454
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: US 09/829,671
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: PCT/US01/40483
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: US 60/197,508
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: US 09/961,721
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: PCT/US01/29904
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: US 60/235,023
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US 10/045,367
; PRIOR FILING DATE: 2001-11-07
; PRIOR APPLICATION NUMBER: US 60/246,561
; PRIOR FILING DATE: 2000-11-07
; PRIOR APPLICATION NUMBER: US 09/801,275
; PRIOR FILING DATE: 2001-03-06
; PRIOR APPLICATION NUMBER: PCT/US01/07074
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: US 60/187,420
; PRIOR FILING DATE: 2000-03-07
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 39
; LENGTH: 2649
; TYPE: DNA

i ORGANISM: Homo sapiens
us-10-170-789-39

Query Match 84.8%; Score 2645.8; DB 6; Length 2649;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2647; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	214	ATGGCAGCAGCAATGGAAACAGAAACAGCTGGGTGTTGAGATATTTGAAACTGCGACTGT	273
Db	1	ATGGCAGCAGCAATGGAAACAGAAACAGCTGGGTGTTGAGATATTTGAAACTGCGACTGT	60
QY	274	GAGGAGAAATTTGAATCACAGGATCGGCCCTAAATTTGGAGCCCTTTTATTTGTTGAGCGGTAT	333
Db	61	GAGGAGAAATTTGAATCACAGGATCGGCCCTAAATTTGGAGCCCTTTTATTTGTTGAGCGGTAT	120
QY	334	TCCTGGAGTCAGCTTAAAGAGCTGCTTGGCATACAGGAAATATCATGGCTACATGATG	393
Db	121	TCCTGGAGTCAGCTTAAAGAGCTGCTTGGCATACAGGAAATATCATGGCTACATGATG	180
QY	394	GCTAAGGACCACATGATTTTCATGTTTGTGAAGGAATGATCCAGATGGACCTCATTTCA	453
Db	181	GCTAAGGACCACATGATTTTCATGTTTGTGAAGGAATGATCCAGATGGACCTCATTTCA	240
QY	454	GACGAAATCTATTAACCTTGGCATGCTCTGGTGAACACAGAGAAATACACTGTTTATTTCT	513
Db	241	GACGAAATCTATTAACCTTGGCATGCTCTGGTGAACACAGAGAAATACACTGTTTATTTCT	300
QY	514	GAATTTCCAAAACATCAATAGAGCAGCTTAAATGCTCTCTTGGAGCCCTTTTGG	573
Db	301	GAATTTCCAAAACATCAATAGAGCAGCTTAAATGCTCTCTTGGAGCCCTTTTGG	360
QY	574	GATCTTTTTCAGGCAACACTGGACTATGGAATGCTTTCTTACGATTTATCACCAGGAAGTGA	633
Db	361	GATCTTTTTCAGGCAACACTGGACTATGGAATGCTTTCTTACGATTTATCACCAGGAAGTGA	420
QY	634	GAAGAAACCCATTTGGAACAGCTCGGAATTTGCTTTTACGATTTATCACCAGGAAGTGA	693
Db	421	GAAGAAACCCATTTGGAACAGCTCGGAATTTGCTTTTACGATTTATCACCAGGAAGTGA	480
QY	694	ACATTTCTGTTTCAAGCCGGTGTGGAATTTATCAAGTAAAGATGGAGGGCCACAAGGA	753
Db	481	ACATTTCTGTTTCAAGCCGGTGTGGAATTTATCAAGTAAAGATGGAGGGCCACAAGGA	540
QY	754	TTTAGGCAACACCTTTAAGCCCAATCTAGTGGAACTAGTTGTCCTCAACATACGGATG	813
Db	541	TTTAGGCAACACCTTTAAGCCCAATCTAGTGGAACTAGTTGTCCTCAACATACGGATG	600
QY	814	GATCCAAATTTATGCCCGCTGATCCAGACTGGATTGCTTTTATACATAGCAACGATAT	873
Db	601	GATCCAAATTTATGCCCGCTGATCCAGACTGGATTGCTTTTATACATAGCAACGATAT	560
QY	874	TGGATATCTAAACATCGTAACACAGAGAAAGAGAGACTCATTTATGTGCACAATGAGCTA	933
Db	661	TGGATATCTAAACATCGTAACACAGAGAAAGAGAGACTCATTTATGTGCACAATGAGCTA	720
QY	934	GCCAAACATGGGAAGATGTCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAGAGNA	993
Db	721	GCCAAACATGGGAAGATGTCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAGAGNA	780
QY	994	TTTGTATAGATATTTCTGGCTATTTGGTGTGTCCTCAAGCTGAAACAACTCCAGTGTGTG	1053
Db	781	TTTGTATAGATATTTCTGGCTATTTGGTGTGTCCTCAAGCTGAAACAACTCCAGTGTGTG	840
QY	1054	AAAAATTTCTAGAAATCTATATGAAGAAATGATGAATCTGAGGTGGAAATTTATTCATGTT	1113
Db	841	AAAAATTTCTAGAAATCTATATGAAGAAATGATGAATCTGAGGTGGAAATTTATTCATGTT	900
QY	1114	ACATCCCTCTATGTTGGAAACAGGAGGCGAGATTCATTCGGTTATCTCTAAAACAGGTACA	1173
Db	901	ACATCCCTCTATGTTGGAAACAGGAGGCGAGATTCATTCGGTTATCTCTAAAACAGGTACA	960
QY	1174	GCAATCTCTAAAGTCACTTTTAAAGTGTACAGAAATATGATGATCTGAGGAAGGATC	1233

Db	961	GCAATCTCTAAAGTCACTTTTAAAGTGTACAGAAATATGATGATGCTCAAGGAAGGATC	1020
QY	1234	ATAGATGTCTATAGATAAGGAACAAATTTCAACCTTTTGTGAGATTCATTTGGAAGAGTTGAA	1293
Db	1021	ATAGATGTCTATAGATAAGGAACAAATTTCAACCTTTTGTGAGATTCATTTGGAAGAGTTGAA	1080
QY	1294	TATATTGCCAGAGCTGGATGGACTCCCTGAGGGAATAATGCTTGGTCCATCTCTACTAGAT	1353
Db	1081	TATATTGCCAGAGCTGGATGGACTCCCTGAGGGAATAATGCTTGGTCCATCTCTACTAGAT	1140
QY	1354	CGCTCCAGAGCTCGCCTACAGATAGTGTGATCTCACCTGAATTTATTTATCTCCAGTAGAA	1413
Db	1141	CGCTCCAGAGCTCGCCTACAGATAGTGTGATCTCACCTGAATTTATTTATCTCCAGTAGAA	1200
QY	1414	GATGATGTTTATGGAAAGCCAGAGACTCAATTGAGTCAAGTGTGCTGATCTGTGAGCCACTA	1473
Db	1201	GATGATGTTTATGGAAAGCCAGAGACTCAATTGAGTCAAGTGTGCTGATCTGTGAGCCACTA	1260
QY	1474	ATTATCTATGAAGAAACACAGACATCTGGATTAATATCCATGACATCTTTTCAATGTTTT	1533
Db	1261	ATTATCTATGAAGAAACACAGACATCTGGATTAATATCCATGACATCTTTTCAATGTTTT	1320
QY	1534	CCCCAAGTCAAGAGAGGAATTTGATTTTATTTTTCCTCTCTGAATGCAAAAACAGGTTTC	1593
Db	1321	CCCCAAGTCAAGAGAGGAATTTGATTTTATTTTTCCTCTCTGAATGCAAAAACAGGTTTC	1380
QY	1594	CGTCAATTTATACAAATTTACATCTATTTTAAAGAAAGCAAAATATAAACATCAAGTGGT	1653
Db	1381	CGTCAATTTATACAAATTTACATCTATTTTAAAGAAAGCAAAATATAAACATCAAGTGGT	1440
QY	1654	GGGCTGCTGCTCCAAAGTGTTCAGAGTTCCTATCAAGAGGAGATAGCAATTTACCAGT	1713
Db	1441	GGGCTGCTGCTCCAAAGTGTTCAGAGTTCCTATCAAGAGGAGATAGCAATTTACCAGT	1500
QY	1714	GGTGAATGGGAGTTCTTGGCGGCATGGATCTAATATCCAAGTTGATGAAGTCAGAGG	1773
Db	1501	GGTGAATGGGAGTTCTTGGCGGCATGGATCTAATATCCAAGTTGATGAAGTCAGAGG	1560
QY	1774	CTGGTATATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACCTGTAGTACGTAGT	1833
Db	1561	CTGGTATATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACCTGTAGTACGTAGT	1620
QY	1834	TACGTAATCTCTGGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACATTTCTGTGTC	1893
Db	1621	TACGTAATCTCTGGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACATTTCTGTGTC	1680
QY	1894	ATCAGTCAAGCTGTGACTTCTTTATAGTAAGTATAGTACCAAGAGAAATCCACACTGT	1953
Db	1681	ATCAGTCAAGCTGTGACTTCTTTATAGTAAGTATAGTACCAAGAGAAATCCACACTGT	1740
QY	1954	GTGTCCTTTTACAAGCTATCAAGTCTCAAGATGACCCAACTTGCAAAAACAAAGGAATTT	2013
Db	1741	GTGTCCTTTTACAAGCTATCAAGTCTCAAGATGACCCAACTTGCAAAAACAAAGGAATTT	1800
QY	2014	TGGGCCCAACATTTTGGATTCAGCAGGTCTCTCTCTGACTATATCTCTCCAGAAAATTTTC	2073
Db	1801	TGGGCCCAACATTTTGGATTCAGCAGGTCTCTCTCTGACTATATCTCTCCAGAAAATTTTC	1860
QY	2074	TCCTTTGAAAGTACTACTGGAATTTACATTTGATGGAGTCTTACAGCCCTCATGATCTA	2133
Db	1861	TCCTTTGAAAGTACTACTGGAATTTACATTTGATGGAGTCTTACAGCCCTCATGATCTA	1920
QY	2134	CAGCCTGGAAAGAAATATCTACTGCTGTGTTTATATATATGTTGTTGCTCAGGTGAGTTG	2193
Db	1921	CAGCCTGGAAAGAAATATCTACTGCTGTGTTTATATATATGTTGTTGCTCAGGTGAGTTG	1980
QY	2194	GTGAATAATTCGGTTTAAAGGAGTCAAGTATTTTCCCTTTGAATACCTTAGCCCTCTTAGT	2253
Db	1981	GTGAATAATTCGGTTTAAAGGAGTCAAGTATTTTCCGCTTTGAATACCTTAGCCCTCTTAGT	2040
QY	2254	TATGTGTTTGTAGTATAGACAAACAGGGGATCTCTGTACCGAGGGCTTAAATTTGAAGGC	2313
Db	2041	TATGTGTTTGTAGTATAGACAAACAGGGGATCTCTGTACCGAGGGCTTAAATTTGAAGGC	2100

Qy	2314	GCCTTTAAATATAAATGGGTCAAAATAGAAATTTGACGATCAGGTGGAAGGACTCCAATAT	2373
Db	2101	GCCTTTAAATATAAATGGGTCAAAATAGAAATTTGACGATCAGGTGGAAGGACTCCAATAT	2160
Qy	2374	CTAGCTCTCGATATGATTTTCATTTAGCTTAGATCGTGTGGGATCCACGGCTGGTCCCTAT	2433
Db	2161	CTAGCTCTCGATATGATTTTCATTTAGCTTAGATCGTGTGGGATCCACGGCTGGTCCCTAT	2220
Qy	2434	GGAGGATACCTCTCCCTGATGGCATTAATGCAGAGTCAAGATCTTTCAGGGTTGCTATT	2493
Db	2221	GGAGGATACCTCTCCCTGATGGCATTAATGCAGAGTCAAGATCTTTCAGGGTTGCTATT	2280
Qy	2494	GCTGGGGCCCCGAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACTTTATATG	2553
Db	2281	GCTGGGGCCCCGAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACTTTATATG	2340
Qy	2554	GGTCAACCTCGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAA	2613
Db	2341	GGTCAACCTCGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAA	2400
Qy	2614	AAGTTCCCTCTGAAACCAAAATCGTTTACTGCTCTTACATGGTTTCTGGATGAGAAATGTC	2673
Db	2401	AAGTTCCCTCTGAAACCAAAATCGTTTACTGCTCTTACATGGTTTCTGGATGAGAAATGTC	2460
Qy	2674	CATTTTGCACATACAGTATATTACTGAGTTTTTTTAGTGAGGCTGGAAAGCCATATGAT	2733
Db	2461	CATTTTGCACATACAGTATATTACTGAGTTTTTTTAGTGAGGCTGGAAAGCCATATGAT	2520
Qy	2734	TTACAGATCTATCTCAGGAGACACAGCATAAAGATTCCCTGAATCGGAGAACATTAT	2793
Db	2521	TTACAGATCTATCTCAGGAGACACAGCATAAAGATTCCCTGAATCGGAGAACATTAT	2580
Qy	2794	GAATGCACTTTTGGCACTACCTCAAGAAACCTTGGATCAGCTATTCTGCTCTTAA	2853
Db	2581	GAATGCACTTTTGGCACTACCTCAAGAAACCTTGGATCAGCTATTCTGCTCTTAA	2640
Qy	2854	GTGATATTA 2862	
Db	2641	GTGATATTA 2649	

RESULT 12
US-09-976-674-2
; Sequence 2, Application US/09976674
; Patent No. US2002015843A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 2671
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-976-674-2

Query Match 84.8%; Score 2644.4; DB 3; Length 2671;
Best Local Similarity 95.8%; Fred. No. 0;
Matches 2648; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy	209	GCAATCGGCGAGCAGCAATGGAACAGACAGCTGGGTGGATATTTGAACTGCGG	268
Db	3	GTACATGGCAGCAGCAATGGAACAGACAGCTGGGTGGATATTTGAACTGCGG	62

Qy	269	ACTGTGAGGAGAAATATTGAATCAGGATCGCCTAAATTTGGAGCCCTTTTATGTTGAGC	328
Db	63	ACTGTGAGGAGAAATATTGAATCAGGATCGCCTAAATTTGGAGCCCTTTTATGTTGAGC	122
Qy	329	GGTATTTCTGGAGTCAGCTTAAAGAGCTGTTGCCGATACCAAGAAATATCATGGCTACA	388
Db	123	GGTATTTCTGGAGTCAGCTTAAAGAGCTGTTGCCGATACCAAGAAATATCATGGCTACA	182
Qy	389	TCATGGCTAAGGCCACCATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGGACCTC	448
Db	183	TCATGGCTAAGGCCACCATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGGACCTC	242
Qy	449	ATTTCAGACAGAAATCTATTACCTTCGCCATGCTGTGTGAGAAACAGAGAAATACTCTGTTTT	508
Db	243	ATTTCAGACAGAAATCTATTACCTTCGCCATGCTGTGTGAGAAACAGAGAAATACTCTGTTTT	302
Qy	509	ATTCTGAAATTTCCAAAACTATCAATAGACAGCAGTCTTAAATGCTCTCTTGGAAAGCCCTC	568
Db	303	ATTCTGAAATTTCCAAAACTATCAATAGACAGCAGTCTTAAATGCTCTCTTGGAAAGCCCTC	362
Qy	569	TTTTGGATCTTTTTTCAGGCAACACTGGACTATGGAATGTATTTCTCGAGAGAAAGAACTAT	628
Db	363	TTTTGGATCTTTTTTCAGGCAACACTGGACTATGGAATGTATTTCTCGAGAGAAAGAACTAT	422
Qy	629	TAAAGAGAAAGAAACCGCATTCGAAACAGTCCGAAATGCTCTTACGATTAACCAAGGAA	688
Db	423	TAAAGAGAAAGAAACCGCATTCGAAACAGTCCGAAATGCTCTTACGATTAACCAAGGAA	482
Qy	689	GTGGAACATTTCTGTTTCAAGCCGGTAGTGGAAATTTATCACGTAAAAGATGGAGGCCAC	748
Db	483	GTGGAACATTTCTGTTTCAAGCCGGTAGTGGAAATTTATCACGTAAAAGATGGAGGCCAC	542
Qy	749	AAGGATTTACGCAACAACTTTTAAAGGCCCAATCTAGTGGAAACTAGTTGTCCCAACATAC	808
Db	543	AAGGATTTACGCAACAACTTTTAAAGGCCCAATCTAGTGGAAACTAGTTGTCCCAACATAC	602
Qy	809	GGATGGATCCAAATATGTCGCCCTGATCCAGATCCGATGCTTTTATACATAGCAAG	868
Db	603	GGATGGATCCAAATATGTCGCCCTGATCCAGATCCGATGCTTTTATACATAGCAAG	662
Qy	869	ATATTTGGATCTTAAACATCGTAAACAGAGAGAAAGGAGACTCATTATGTGCACATG	928
Db	663	ATATTTGGATCTTAAACATCGTAAACAGAGAGAAAGGAGACTCATTATGTGCACATG	722
Qy	929	AGCTAGCCAAATGGAAGAGATCCAGATCAGCTGGAGTCGCTAGCTTTGTTTCTCAAG	988
Db	723	AGCTAGCCAAATGGAAGAGATCCAGATCAGCTGGAGTCGCTAGCTTTGTTTCTCAAG	782
Qy	989	AAGAAATTTGATAGATATTCTGGCTATTGGTGGTGTCCAAAAGCTGAAACAACTCCAGTG	1048
Db	783	AAGAAATTTGATAGATATTCTGGCTATTGGTGGTGTCCAAAAGCTGAAACAACTCCAGTG	842
Qy	1049	GTGGTAAAAATCTTAGAATTTCTATGAAGAAATGATGAATCTGAGGTGGAATTTATTC	1108
Db	843	GTGGTAAAAATCTTAGAATTTCTATGAAGAAATGATGAATCTGAGGTGGAATTTATTC	902
Qy	1109	ATGTTTACATCCCTATGTTGGAACAAAGGAGGCGAGATTCATTCCGTTATCTTAAACAG	1168
Db	903	ATGTTTACATCCCTATGTTGGAACAAAGGAGGCGAGATTCATTCCGTTATCTTAAACAG	962
Qy	1169	GTACAGCAAACTCTAAAGTCACTTTTAAAGATGTGAGAAATGAATGATGCTGGAAGGAA	1228
Db	963	GTACAGCAAACTCTAAAGTCACTTTTAAAGATGTGAGAAATGAATGATGCTGGAAGGAA	1022
Qy	1229	GGATCATAGATGTCATAGATGAAGAACTAAATCNAACCTTTTGAGATCTATTTGAGGAG	1288
Db	1023	GGATCATAGATGTCATAGATGAAGAACTAAATCNAACCTTTTGAGATCTATTTGAGGAG	1082
Qy	1289	TTGAATATATTTGCCAGAGCTGGATGGAATCTGAGGGGAAATATGCTGCTCATCTAC	1348
Db	1083	TTGAATATATTTGCCAGAGCTGGATGGAATCTGAGGGGAAATATGCTGCTCATCTAC	1142
Qy	1349	TAGATCGCTCCAGACTCGCCTACAGATAGTGTGATCTCACCTGAATTTATTTATCCCAG	1408

```
Db 1143 TAGATCGCTCCAGACTCGCTCGAGATAGTGTGATCTCACTGATTTATTTATCCAG 1202
Qy 1409 TAGAAGATGATGTTATGGAAGGAGAGACTCAATTGAGTCAGTGCCTGATTTCTGTGACGC 1468
Db 1203 TAGAAGATGATGTTATGGAAGGAGAGACTCAATTGAGTCAGTGCCTGATTTCTGTGACGC 1262
Qy 1469 CACTAATTTATCTATGAGAAACAACAGACATCTGGATAAATATCAATGATCATCTTTTCATG 1528
Db 1263 CACTAATTTATCTATGAGAAACAACAGACATCTGGATAAATATCAATGATCATCTTTTCATG 1322
Qy 1529 TTTTTCCTCCAAAGTCAGAGAGAAATTCAGTTTATTTTGGCTCTGAAATCAAAACAG 1588
Db 1323 TTTTTCCTCCAAAGTCAGAGAGAAATTCAGTTTATTTTGGCTCTGAAATCAAAACAG 1382
Qy 1589 GTTTCCTGTCATTTATACAAATATCATCTATTTTAAAGGAAAGCAAAATATAACGATCCA 1648
Db 1383 GTTTCCTGTCATTTATACAAATATCATCTATTTTAAAGGAAAGCAAAATATAACGATCCA 1442
Qy 1649 GTGGTGGCTGCTGCTCCAAAGTGAATTTCAAGTGTCTATCAAGAGGAGATAGCAATTA 1708
Db 1443 GTGGTGGCTGCTGCTCCAAAGTGAATTTCAAGTGTCTATCAAGAGGAGATAGCAATTA 1502
Qy 1709 CCAGTGGTGAATGGAGTCTTTCGCGGCATGGATCTAATATCCAAAGTTGATGAGTCA 1768
Db 1503 CCAGTGGTGAATGGAGTCTTTCGCGGCATGGATCTAATATCCAAAGTTGATGAGTCA 1562
Qy 1769 GAAGCTGCTGATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTACGTAG 1828
Db 1563 GAAGCTGCTGATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTACGTAG 1622
Qy 1829 TCAGTTAGCTAAATCTCGAGAGGTGAACAAGGTGACTGACCGTGGCTACTACATCTTT 1888
Db 1623 TCAGTTAGCTAAATCTCGAGAGGTGAACAAGGTGACTGACCGTGGCTACTACATCTTT 1682
Qy 1889 GTGCAATCAGTCAGCACTGTGACTCTTTTATAGTAAGTATAGTAACCAAGAGATCCAC 1948
Db 1683 GTGCAATCAGTCAGCACTGTGACTCTTTTATAGTAAGTATAGTAACCAAGAGATCCAC 1742
Qy 1949 ACTGTGTGCTCCTTTACAAAGCTATCAAGTCTGAAAGTACCAACTTGCMAAACAAAGG 2008
Db 1743 ACTGTGTGCTCCTTTACAAAGCTATCAAGTCTGAAAGTACCAACTTGCMAAACAAAGG 1802
Qy 2009 AATTTTGGGCCACCAATTTGGATTTCAGCAGGTCTCTTCTGACTACTACTCTCCAGAAA 2068
Db 1803 AATTTTGGGCCACCAATTTGGATTTCAGCAGGTCTCTTCTGACTACTACTCTCCAGAAA 1862
Qy 2069 TTTTCTCTTTTGAAGTACTACTGGAATTTACATTTGATGGGATGCTCTACAGCCTCATG 2128
Db 1863 TTTTCTCTTTTGAAGTACTACTGGAATTTACATTTGATGGGATGCTCTACAGCCTCATG 1922
Qy 2129 ATCTACAGCTGGAAAGAAATATCCTACTGTGCTGTTTATATATGTTGGTCTTCAGGTGC 2188
Db 1923 ATCTACAGCTGGAAAGAAATATCCTACTGTGCTGTTTATATATGTTGGTCTTCAGGTGC 1982
Qy 2189 AGTTGTGTAATTAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTAGCCTCTC 2248
Db 1983 AGTTGTGTAATTAATCGAATTTAAAGAGTCAAGTATTTCCGCTTGAATACCTAGCCTCTC 2042
Qy 2249 TAGGTTATGCTGTTGATGATAGACAAACAGGGGATCTGTCAACCGAGGCTTAAATTTG 2308
Db 2043 TAGGTTATGCTGTTGATGATAGACAAACAGGGGATCTGTCAACCGAGGCTTAAATTTG 2102
Qy 2309 AAGGGCTCTTTAAATATATAAATGGTCAAAATAGAAATTCAGATTCAGGTGGAAGGACTCC 2368
Db 2103 AAGGGCTCTTTAAATATATAAATGGTCAAAATAGAAATTCAGATTCAGGTGGAAGGACTCC 2162
Qy 2369 AATATCTAGCTTCTCGATATGATTTTCAATTGACTTAGATCGTGTGGGCATCCACGGCTGGT 2428
Db 2163 AATATCTAGCTTCTCGATATGATTTTCAATTGACTTAGATCGTGTGGGCATCCACGGCTGGT 2222
Qy 2429 CCTATGGAGGATACCTCTCCCTGATGGCATTTAATGCAGAGTCAAGATATCTTCAGGGTTG 2488
```

```
Db 2223 CCTATGGAGGATACCTCTCCCTGATGGCATTTAATGACAGGTGATATCTTCAGGGTTG 2282
Qy 2489 CTATTTGCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACGTT 2548
Db 2283 CTATTTGCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACGTT 2342
Qy 2549 ATATGGGTCACTCCCTGACCAAGTAAAGAGGCTATTACTTTAGGATCTGTGGCCATGCAAG 2608
Db 2343 ATATGGGTCACTCCCTGACCAAGTAAAGAGGCTATTACTTTAGGATCTGTGGCCATGCAAG 2402
Qy 2609 CAGAAAGTTCCCTCTCTGAAACCAATCGTTTACTGTCTTTACATGGTTTCTTGATGAGA 2668
Db 2403 CAGAAAGTTCCCTCTCTGAAACCAATCGTTTACTGTCTTTACATGGTTTCTTGATGAGA 2462
Qy 2669 ATGTCCATTTTGCATACATACAGTATATTACTAGTGTTTTATGAGGCTGGAAGCCAT 2728
Db 2463 ATGTCCATTTTGCATACATACAGTATATTACTAGTGTTTTATGAGGCTGGAAGCCAT 2522
Qy 2729 ATGATTTACAGATCTATCTCTCAGGAGAGACACAGCATAAAGTTCCTGAAATCGGGAGAAC 2788
Db 2523 ATGATTTACAGATCTATCTCTCAGGAGAGACACAGCATAAAGTTCCTGAAATCGGGAGAAC 2582
Qy 2789 ATTATGAATCGATCTTTTTCGACTACCTTCAAGAAACCTTTGGATCAGTATTGCTGCTC 2848
Db 2583 ATTATGAATCGATCTTTTTCGACTACCTTCAAGAAACCTTTGGATCAGTATTGCTGCTC 2642
Qy 2849 TAAAGTGATATAA 2862
Db 2643 TAAAGTGATATAA 2656
```

RESULT 13

```
US-10-982-512-2
; Sequence 2, Application US/10982512
; Publication No. US20050059081A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/10/982,512
; CURRENT FILING DATE: 2004-11-05
; PRIOR APPLICATION NUMBER: US/09/976,674
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 2
; LENGTH: 2671
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-982-512-2
```

```
Query Match 84.8%; Score 2644.4; DB 9; Length 2671;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2648; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 209 GCACATGGCAGCAGCAATGGAAAACAGAACACAGTGGTGTTCAGATATTTGAAATCGCG 268
Db 3 GTACCATGGCAGCAGCAATGGAAAACAGAACACAGTGGTGTTCAGATATTTGAAATCGCG 62

Qy 269 ACTGTGAGGAGATATTGAAATCAGAGTCGGCTTAAATTTGAGGCTTTTATGTTGAGC 328
Db 63 ACTGTGAGGAGATATTGAAATCAGAGTCGGCTTAAATTTGAGGCTTTTATGTTGAGC 122

Qy 329 GGTATTCTCGAGTCAGCTTAAAGCTCTTTCCGATACCAAGAAATATCATGGCTACA 388
Db 123 GGTATTCTCGAGTCAGCTTAAAGCTCTTTCCGATACCAAGAAATATCATGGCTACA 182

Qy 389 TGATGGCTAAGGCCACCATGATTTTCAATGTTGTGAAGGAGATGATCCAGATGGACCTC 448
```

Db	183	TGATGGCTAAGGCACCACTGATTTCTATGTTGTGAAGAGGAATGATCCAGATGACCTC	242	Db	1263	CACTAATTATCTATGAAGAAACAACAGACATCTCGATAAATATATCCATGACATCTTTTCATG	1322
Qy	449	ATTGACAGAGAAATCTATTACCTTCCCATGTCCTGGTGAGAACAGAGAAATATACACTGTTTT	508	Qy	1529	TTTTTCCCAAGTCAACGAAGAGAAATTTGAGTTTATTTTGGCTCTGTAATGCAAAACAG	1588
Db	243	ATTGACAGAGAAATCTATTACCTTCCCATGTCCTGGTGAGAACAGAGAAATATACACTGTTTT	302	Db	1323	TTTTTCCCAAGTCAACGAAGAGAAATTTGAGTTTATTTTGGCTCTGTAATGCAAAACAG	1382
Qy	509	ATTCTGAAATTTCCCAAAACTATCAATAGAGACGAGCTTTAAATGCTCTCTTGGAGGCCTC	568	Qy	1589	GTTTCGGTCATTTATACAAAATTTACATCTATTTTAAAGGAAGCAATATAAAGCATCCA	1648
Db	303	ATTCTGAAATTTCCCAAAACTATCAATAGAGACGAGCTTTAAATGCTCTCTTGGAGGCCTC	362	Db	1383	GTTTCGGTCATTTATACAAAATTTACATCTATTTTAAAGGAAGCAATATAAAGCATCCA	1442
Qy	569	TTTTGGATCTTTTTCAGSCAACTGGAATATGGAATGTTATCTCGAGAAGAGAACTAT	628	Qy	1649	GTGGTGGGCTGCTCTCCAAAGTGATTTCAAGTGCTCTATCAAGAGAGAGATAGCAATTA	1708
Db	363	TTTTGGATCTTTTTCAGSCAACTGGAATATGGAATGTTATCTCGAGAAGAGAACTAT	422	Db	1443	GTGGTGGGCTGCTCTCCAAAGTGATTTCAAGTGCTCTATCAAGAGAGAGATAGCAATTA	1502
Qy	629	TAAAGAGAAAGAAACGCAATTCGGAACAGTCGGAATTTGCTTTACAGATTTATCAACCAAGAA	688	Qy	1709	CCAGTGGTGAATGGGAAGTTCTTTGGCCGGCATGGATCTAATATCAAGTTGATGAAGTCA	1768
Db	423	TAAAGAGAAAGAAACGCAATTCGGAACAGTCGGAATTTGCTTTACAGATTTATCAACCAAGAA	482	Db	1503	CCAGTGGTGAATGGGAAGTTCTTTGGCCGGCATGGATCTAATATCAAGTTGATGAAGTCA	1562
Qy	689	GTGGAACTTTCTGTTTTCAAGCCGGTAGTGGAAATTTATCAAGTAAAAAGATGGAGGCCAC	748	Qy	1769	GAAGGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACCTGTACGTAG	1828
Db	483	GTGGAACTTTCTGTTTTCAAGCCGGTAGTGGAAATTTATCAAGTAAAAAGATGGAGGCCAC	542	Db	1563	GAAGGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACCTGTACGTAG	1622
Qy	749	RAGGATTTACGCAACACCTTTAAGGCCCAATCTAGTGGAAACTAGTGTCTCCCAACATAC	808	Qy	1829	TCAGTTTACGTAAATCCTCGAGAGGTGACAAAGCTGACTGACCGTGGCTACTCACATTTCTT	1888
Db	543	RAGGATTTACGCAACACCTTTAAGGCCCAATCTAGTGGAAACTAGTGTCTCCCAACATAC	602	Db	1623	TCAGTTTACGTAAATCCTCGAGAGGTGACAAAGCTGACTGACCGTGGCTACTCACATTTCTT	1682
Qy	809	GGATGGATCCAAATTTATGCCCGCTGATCCAGACTGGATTTGCTTTTATACATAGCAAG	868	Qy	1889	GCTGCATCAGTCAGCACTGTGACTTTCTTTATAAGTAAAGTATAGTAACCAAGAAATCCAC	1948
Db	603	GGATGGATCCAAATTTATGCCCGCTGATCCAGACTGGATTTGCTTTTATACATAGCAAG	662	Db	1683	GCTGCATCAGTCAGCACTGTGACTTTCTTTATAAGTAAAGTATAGTAACCAAGAAATCCAC	1742
Qy	869	ATATTTGGATTTCTAACATCTGTAAACAGAGAGAAAGAGAGACTCACCTTATGTGCAATG	928	Qy	1949	ACTGTGTGCTTTTACAGCTATCAGCTATCAGCTCTGAGATGACCCAACTTTGCAAAACAAAGG	2008
Db	663	ATATTTGGATTTCTAACATCTGTAAACAGAGAGAAAGAGAGACTCACCTTATGTGCAATG	722	Db	1743	ACTGTGTGCTTTTACAGCTATCAGCTCTGAGATGACCCAACTTTGCAAAACAAAGG	1802
Qy	929	AGCTAGCAACATGGAAGAAGATGCCAGATCAGCTGGAGTCGTACTCTTTGTTCTCCAAG	988	Qy	2009	AATTTTGGGCAACCAATTTTGGATTCAGCAGGTCTCTCTCTGACTATATCTCTCCAGAAA	2068
Db	723	AGCTAGCAACATGGAAGAAGATGCCAGATCAGCTGGAGTCGTACTCTTTGTTCTCCAAG	782	Db	1803	AATTTTGGGCAACCAATTTTGGATTCAGCAGGTCTCTCTCTGACTATATCTCTCCAGAAA	1862
Qy	989	AAGAAATTTGATAGATATTCTGGCTATTGGTGGTGTCCAAAGCTGAAACAACTCCCAAGT	1048	Qy	2069	TTTTCTTTTTGAAAGTACTACTGATTTTACATTTGATGGGATGCTCTCAAGGCTCATG	2128
Db	783	AAGAAATTTGATAGATATTCTGGCTATTGGTGGTGTCCAAAGCTGAAACAACTCCCAAGT	842	Db	1863	TTTTCTTTTTGAAAGTACTACTGATTTTACATTTGATGGGATGCTCTCAAGGCTCATG	1922
Qy	1049	GTGGTAAAAATTTCTAGAAATTTCTATGAAGAAATGATGAATCTGAGGTGGAAATTTATTC	1108	Qy	2129	ATCTACAGCTCGAAAGAAATATCCTACTGTCTGTTTATATATATGTTGGTCTCAGGTGC	2188
Db	843	GTGGTAAAAATTTCTAGAAATTTCTATGAAGAAATGATGAATCTGAGGTGGAAATTTATTC	902	Db	1923	ATCTACAGCTCGAAAGAAATATCCTACTGTCTGTTTATATATATGTTGGTCTCAGGTGC	1982
Qy	1109	ATGTTTACATCCCTATGTTGAAACAAGGAGGCGAGATTCATTCCTGTTATCTTAAACAG	1168	Qy	2189	AGTTGGTGAATTAATCGATTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTAGCCTCTC	2248
Db	903	ATGTTTACATCCCTATGTTGAAACAAGGAGGCGAGATTCATTCCTGTTATCTTAAACAG	962	Db	1983	AGTTGGTGAATTAATCGATTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTAGCCTCTC	2042
Qy	1169	GTACAGCAATCCCTAAAGTCACTTTTAAAGTGTGAGAAATAATGATTTGATGCTGAGGAA	1228	Qy	2249	TAGGTTATCTGTGTTGTAGTGATAGACAAACAGGGATCCTGTCAACCGAGGGCTTAAATTTG	2308
Db	963	GTACAGCAATCCCTAAAGTCACTTTTAAAGTGTGAGAAATAATGATTTGATGCTGAGGNA	1022	Db	2043	TAGGTTATCTGTGTTGTAGTGATAGACAAACAGGGATCCTGTCAACCGAGGGCTTAAATTTG	2102
Qy	1229	GGATCATAGATGTCATAGATAAGGAATTAATTCACCTTTTGAATTTCTATTTGAGGAG	1288	Qy	2309	AAGGCGCTTTTAAATATAAATGGGTCAAAATAGAAATTTGACGATCAGGTGGAAGGACTCC	2368
Db	1023	GGATCATAGATGTCATAGATAAGGAATTAATTCACCTTTTGAATTTCTATTTGAGGAG	1082	Db	2103	AAGGCGCTTTTAAATATAAATGGGTCAAAATAGAAATTTGACGATCAGGTGGAAGGACTCC	2162
Qy	1289	TTGAAATATATTGCGAGCTGGATGGAATCTCGAGGGAATAATGCTGGTCCATCTTAC	1348	Qy	2369	AATATCTAGCTTCTCGATATGATTTTCAATTTAGCTTAGATCTGTTGGGCATCCACGGCTGTT	2428
Db	1083	TTGAAATATATTGCGAGCTGGATGGAATCTCGAGGGAATAATGCTGGTCCATCTTAC	1142	Db	2163	AATATCTAGCTTCTCGATATGATTTTCAATTTAGCTTAGATCTGTTGGGCATCCACGGCTGTT	2222
Qy	1349	TAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACTGAAATTTATTTATCCAG	1408	Qy	2429	CCTATGGAAGATACCTTCCCTGATGGCAATTAATGACAGGTGAGATATCTTCAAGGTTG	2488
Db	1143	TAGATCGCTCCAGACTCGCTCCAGATAGTGTGATCTCACTGAAATTTATTTATCCAG	1202	Db	2223	CCTATGGAAGATACCTTCCCTGATGGCAATTAATGACAGGTGAGATATCTTCAAGGTTG	2282
Qy	1409	TAGAAAGATGATTTATGAAAGGAGAGACTCATTTGAGTCAAGTCTGTTGAGCGC	1468	Qy	2489	CTATTGCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGACGTT	2548
Db	1203	TAGAAAGATGATTTATGAAAGGAGAGACTCATTTGAGTCAAGTCTGTTGAGCGC	1262	Db	2283	CTATTGCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGACGTT	2342
Qy	1469	CACTAATTATCTATGAAGAAACAACAGACATCTGGATAAATATCCATGACATCTTTTCATG	1528	Qy	2549	ATATGGGTCACTCGACCAAGATGAACAGGCTATTACTTAGGATCTGTGGCCATGCAAG	2608
		CACTAATTATCTATGAAGAAACAACAGACATCTGGATAAATATCCATGACATCTTTTCATG		Db	2343	ATATGGGTCACTCGACCAAGATGAACAGGCTATTACTTAGGATCTGTGGCCATGCAAG	2402

QY 2609 CAGAAAAGTTCCCTCTGAACCAAAATCGTTTACTGCTCTTACATGGTTTCTCGATGAGA 2668
Db 2403 CAGAAAAGTTCCCTCTGAACCAAAATCGTTTACTGCTCTTACATGGTTTCTCGATGAGA 2462
QY 2669 ATGTCCATTTTGCACATACACAGTATATTAATGAGTTTATGAGGGCTGGAAGCCAT 2728
Db 2463 ATGTCCATTTTGCACATACACAGTATATTAATGAGTTTATGAGGGCTGGAAGCCAT 2522
QY 2729 ATGATTTTACAGATCTATCTCTCAGGAGAGACACAGATTAAGAGTTTCTGAAATCGGGAGAAC 2788
Db 2523 ATGATTTTACAGATCTATCTCTCAGGAGAGACACAGATTAAGAGTTTCTGAAATCGGGAGAAC 2582
QY 2789 ATTATGAATCGCATCTTTTGCACACTACCTTCAAGAAAAACCTTGGATCAGCTATTGCTGCTC 2848
Db 2583 ATTATGAATCGCATCTTTTGCACACTACCTTCAAGAAAAACCTTGGATCAGCTATTGCTGCTC 2642
QY 2849 TAAAAGTGATATAA 2862
Db 2643 TAAAAGTGATATGA 2656

RESULT 14

US-10-054-776-1
; Sequence 1, Application US/10054776
; Publication No. US20030165818A1
; GENERAL INFORMATION:
; APPLICANT: Mark Robert Edbrooke
; APPLICANT: Alan Peter Lewis
; TITLE OF INVENTION: NOVEL PROTEIN
; FILE REFERENCE: QG1042US
; CURRENT APPLICATION NUMBER: US/10/054,776
; CURRENT FILING DATE: 2002-01-23
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 2649
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2649)
US-10-054-776-1

Query Match 84.7%; Score 2644.2; DB 6; Length 2649;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2646; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 214 ATGCGCAGCAGCAATGGAAAACAGAAACAGCTGGGTGTGAGATATTTGAAACTGCGGACTGT 273
Db 1 ATGCGCAGCAGCAATGGAAAACAGAAACAGCTGGGTGTGAGATATTTGAAACTGCGGACTGT 60
QY 274 GAGGAGAAATTTGAATCAGAGATCGGCTAAATTTGGAGCCCTTTTATGTTGACGGGTAT 333
Db 61 GAGGAGAAATTTGAATCAGAGATCGGCTAAATTTGGAGCCCTTTTATGTTGACGGGTAT 120
QY 334 TCCTGGAGTCAGCTTAAAGCTGCTTGGATCCGATACCAAGAAATATCATGGCTACATCATG 393
Db 121 TCCTGGAGTCAGCTTAAAGCTGCTTGGATCCGATACCAAGAAATATCATGGCTACATCATG 180
QY 394 GCTAAGCACCACATGATTTTATGTTGTAAGAGGAATGATCCAGATGGACCTCATTTCA 453
Db 181 GCTAAGGCAACCATGATTTTATGTTGTAAGAGGAATGATCCAGATGGACCTCATTTCA 240
QY 454 GACAGAAATCTATTAACCTTGCCATGCTCTGGTGAGAACAGAGAAATPACACTGTTTATTTCT 513
Db 241 GACAGAAATCTATTAACCTTGCCATGCTCTGGTGAGAACAGAGAAATPACACTGTTTATTTCT 300
QY 514 GAAATTCGAAAACATATCAATAGACAGCAGCTTAAAGCTCTCTTGGAGCCCTTTTGG 573
Db 301 GAAATTCGAAAACATATCAATAGACAGCAGCTTAAAGCTCTCTTGGAGCCCTTTTGG 360
QY 574 GATCTTTTTCAGGCAACACTGGACTATGGAATGTATTCTTCGAGAGAGAAAGAACTATTAGA 633

Db 361 GATCTTTTTCAGGCAACACTGGACTATGGAATGTATTCTCGAGAGAGAAAGAACTATTAGA 420
QY 634 GAAAGAAAAACGCAATTTGAAACAGTCGGAATTTGCTTTACGATTTATCAACAAAGGAAGTGA 693
Db 421 GAAAGAAAAACGCAATTTGAAACAGTCGGAATTTGCTTTACGATTTATCAACAAAGGAAGTGA 480
QY 694 ACATTTCTGTTTCAAGCCGCTAGTGAATTTATACGCTAAAGAGTGGAGGGCCCAAGGA 753
Db 481 ACATTTCTGTTTCAAGCCGCTAGTGAATTTATCAACGCTAAAGAGTGGAGGGCCCAAGGA 540
QY 754 TTTACGCAACCAACCTTTAAAGGCCCAATCTAGTGGAAAACCTAGTGTCCCAACATACGGATG 813
Db 541 TTTACGCAACCAACCTTTAAAGGCCCAATCTAGTGGAAAACCTAGTGTCCCAACATACGGATG 600
QY 814 GATCCAAAATTTATGCCCTGCTGATCCAGACTGGATTTCTTTATACATAGCAACGATATT 873
Db 601 GATCCAAAATTTATGCCCTGCTGATCCAGACTGGATTTCTTTATACATAGCAACGATATT 660
QY 874 TGGATATCTAAACATCTGTAACCAAGAGAGAAAGAGAGACTCACTTATGTGCACAAATGAGCTA 933
Db 661 TGGATATCTAAACATCTGTAACCAAGAGAGAAAGAGAGACTCACTTATGTGCACAAATGAGCTA 720
QY 934 GCCAATCTGAAGAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAAGAGAA 993
Db 721 GCCAATCTGAAGAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAAGAGAA 780
QY 994 TTTGATAGATATTTCTGGCTATTTGGTGGTGTCCAAAAGCTGAAACAACTCCCAAGTGGTGT 1053
Db 781 TTTGATAGATATTTCTGGCTATTTGGTGGTGTCCAAAAGCTGAAACAACTCCCAAGTGGTGT 840
QY 1054 AAAATCTTTAGAAATTTCTATATGAAGAAAATGATGAATCTGAGGTGGGAAAATTTATCATGTT 1113
Db 841 AAAATCTTTAGAAATTTCTATATGAAGAAAATGATGAATCTGAGGTGGGAAAATTTATCATGTT 900
QY 1114 ACATCCCTATGTTGGAAAACAGAGGGGAGAGATTCATTTCCGTTATCTTAAACAGAGTACA 1173
Db 901 ACATCCCTATGTTGGAAAACAGAGGGGAGAGATTCATTTCCGTTATCTTAAACAGAGTACA 960
QY 1174 GCAATCTTAAAGTCACTTTTAAAGATGTCAGAAAATTAATGATGCTGTAAGAGGAGATC 1233
Db 961 GCAATCTTAAAGTCACTTTTAAAGATGTCAGAAAATTAATGATGCTGTAAGAGGAGATC 1020
QY 1234 ATAGATCTATAGATAAGGAATTAATTAACCTTTTGAATTTCTATTTGAAAGGAGTTGAA 1293
Db 1021 ATAGATCTATAGATAAGGAATTAATTAACCTTTTGAATTTCTATTTGAAAGGAGTTGAA 1080
QY 1294 TATATCCAGAGCTGGATGGACTCTGAGGAAAATATGCTGGTCCATCTACTAGAT 1353
Db 1081 TATATCCAGAGCTGGATGGACTCTGAGGAAAATATGCTGGTCCATCTACTAGAT 1140
QY 1354 CGCTCCAGAGCTCGCTTACAGATAGTGTGTTGATCTCACTGAAATTTATTTATCCAGTAGAA 1413
Db 1141 CGCTCCAGAGCTCGCTTACAGATAGTGTGTTGATCTCACTGAAATTTATTTATCCAGTAGAA 1200
QY 1414 GATGATGTTATGGAAGGAGAGACTCAATTTGAGTCAGTCCCTGATTTGTGAAGCCACTA 1473
Db 1201 GATGATGTTATGGAAGGAGAGACTCAATTTGAGTCAGTCCCTGATTTGTGAAGCCACTA 1260
QY 1474 ATTATCTATGAAGAAACACAGACATCTCGATTAATATCCATGACATCTTTTCATGTTTTC 1533
Db 1261 ATTATCTATGAAGAAACACAGACATCTCGATTAATATCCATGACATCTTTTCATGTTTTC 1320
QY 1534 CCCCAAGTCCAGAGGAGAAATTTGAGTTTATTTTGGCTCTGAAATGCAAAACAGGTTTC 1593
Db 1321 CCCCAAGTCCAGAGGAGAAATTTGAGTTTATTTTGGCTCTGAAATGCAAAACAGGTTTC 1380
QY 1594 CGTCAATTTATACAAAATTAATCTATTTTAAAGGAGAAAGCAAAATATAAAGATCCAGTGGT 1653
Db 1381 CGTCAATTTATACAAAATTAATCTATTTTAAAGGAGAAAGCAAAATATAAAGATCCAGTGGT 1440
QY 1654 GGGTGGCTGCTCAAGTGAATTTCAAGTGTCTTATCAAGAGGAGATAGCAATTTACCAAGT 1713

1441 GGGCTGCTCTCCAAAGTGATTTCAAGTGTCTCTATCAAAAGAGGAGATAGCAATATACCAAGT 1500
Db
1714 GGTGAATGGGAAGTTCTTTGGCGGCATGGATCTAAATATCAAGTTTGAATGAAGTCAAGG 1773
Qy
1501 GGTGAATGGGAAGTTCTTTGGCGGCATGGATCTAAATATCAAGTTTGAATGAAGTCAAGG 1560
Db
1774 GTGGTATATTTGAAGGCAACAAGACTCCCTTTTAGAGCATCACTGTACGTAGTCAGT 1833
Qy
1561 CTGGTATATTTGAAGGCAACAAGACTCCCTTTTAGAGCATCACTGTACGTAGTCAGT 1620
Db
1834 TACGTAAATCTGGAGAGGTGACAAAGCTGACTGACGGTGGTACTCACATCTTTGCTGC 1893
Qy
1621 TACGTAAATCTGGAGAGGTGACAAAGCTGACTGACGGTGGTACTCACATCTTTGCTGC 1680
Db
1894 ATCAGTCAAGCTGTGACTCTTTTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 1953
Qy
1681 ATCAGTCAAGCTGTGACTCTTTTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 1740
Db
1954 GTGTCCTTTTAAAGCTATCAAGTCTTGAAGATGACCCAACTTTGCAAAACAAGGAATTT 2013
Qy
1741 GTGTCCTTTTAAAGCTATCAAGTCTTGAAGATGACCCAACTTTGCAAAACAAGGAATTT 1800
Db
2014 TGGGGCACCATTGTTGGATTCAGCAGGTCTCTCTTCTGACTATCTCTCCAGAAATTTTC 2073
Qy
1801 TGGGGCACCATTGTTGGATTCAGCAGGTCTCTCTTCTGACTATCTCTCCAGAAATTTTC 1860
Db
2074 TCTTTTGAAGTACTACTGGATTTTACATTTGTATGGGATGCTCTCAAGGCTCATGATCTA 2133
Qy
1861 TCTTTTGAAGTACTACTGGATTTTACATTTGTATGGGATGCTCTCAAGGCTCATGATCTA 1920
Db
2134 CAGCTCGAAGAAATATCTTACTGTGCTGTTTATATATGTTGGTCTCTCAGGTGCAAGTTG 2193
Qy
1921 CAGCTCGAAGAAATATCTTACTGTGCTGTTTATATATGTTGGTCTCTCAGGTGCAAGTTG 1980
Db
2194 GTGAAATATCGGTTTAAAGGATCAAGTATTTTCCGCTTGAATACCCCTAGCTCTCTAGGT 2253
Qy
1981 GTGAAATATCGGTTTAAAGGATCAAGTATTTTCCGCTTGAATACCCCTAGCTCTCTAGGT 2040
Db
2254 TATGCTGTTGTAGTATAGCAACAAGGGATCCCTGCTCAAGGGGCTTAAATTTGAAGGC 2313
Qy
2041 TATGCTGTTGTAGTATAGCAACAAGGGATCCCTGCTCAAGGGGCTTAAATTTGAAGGC 2100
Db
2314 GCCTTTAAATATAAATGGGTCMAATAGAAATTCAGCATCAGGTGGAAGGACTCCAAATAT 2373
Qy
2101 GCCTTTAAATATAAATGGGTCMAATAGAAATTCAGCATCAGGTGGAAGGACTCCAAATAT 2160
Db
2374 CTAGCTTCTCGATATGATTTCACTTGAATATGACAGAGGTCAAGATCTTTCAAGGCTGCTAT 2433
Qy
2161 CTAGCTTCTCGATATGATTTCACTTGAATATGACAGAGGTCAAGATCTTTCAAGGCTGCTAT 2220
Db
2434 GGAGGATACCTCTCCCTGATGGCATTAATGACAGAGGTCAAGATCTTTCAAGGCTGCTAT 2493
Qy
2221 GGAGGATACCTCTCCCTGATGGCATTAATGACAGAGGTCAAGATCTTTCAAGGCTGCTAT 2280
Db
2494 GCTGGGGCCCGCAGTCACTCTGTGGATCTTTCTATGATACAGGATACACGGAACTGTATATG 2553
Qy
2281 GCTGGGGCCCGCAGTCACTCTGTGGATCTTTCTATGATACAGGATACACGGAACTGTATATG 2340
Db
2554 GGTCACTGACACAGAAATGAACAGGGCTATTTACTAGGATCTGTGGCATGCAAGCAGAA 2613
Qy
2341 GGTCACTGACACAGAAATGAACAGGGCTATTTACTAGGATCTGTGGCATGCAAGCAGAA 2400
Db
2614 AAGTTCCTCTCAACCAATCGTTTACTGCTCTTACATGTTTCTCTGGATGAGAATGTC 2673
Qy
2401 AAGTTCCTCTCAACCAATCGTTTACTGCTCTTACATGTTTCTCTGGATGAGAATGTC 2460
Db
2674 CATTTTGCATATCCAGTATATTTACTGAGTTTCTTAAAGTGGGCTGGAAAGCCATATGAT 2733
Qy
2461 CATTTTGCATATCCAGTATATTTACTGAGTTTCTTAAAGTGGGCTGGAAAGCCATATGAT 2520
Db
2734 TTACAGATCTATCTCCAGGAGACACAGCATTAAGGTTCTGAAATCTGGAGACATTTAT 2793
Qy
2521 TTACAGATCTATCTCCAGGAGACACAGCATTAAGGTTCTGAAATCTGGAGACATTTAT 2580
Db

2794 GAACCTGCAATCTTTTGGCACTACCTTCAAGAAAACCTTGGATCAGCTATTCCTCTATAAA 2853
Qy
2581 GAACCTGCAATCTTTTGGCACTACCTTCAAGAAAACCTTGGATCAGCTATTCCTCTATAAA 2640
Db
2854 GTGATATAA 2862
Qy
2641 GTGATATAA 2649
Db
RESULT 15
US-10-956-157-2177
; Sequence 2177, Application US/10956157
; Publication No. US20050118625A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
; TITLE OF INVENTION: HUMAN OSTEOARTHRITIS AND HUMAN PROTEASES
; FILE REFERENCE: 031896-043000 (AM 101081)
; CURRENT APPLICATION NUMBER: US/10/956,157
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 319805
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2177
; LENGTH: 2830
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-956-157-2177
Query Match 80.4%; Score 2508.4; DB 9; Length 2830;
Best Local Similarity 90.4%; Pred. No. 0;
Matches 2819; Conservative 0; Mismatches 1; Indels 300; Gaps 1;
Qy 1 AAGTGCTAAAGCTCCGAGGCCAAGCGCGCTGCTACTGCGCGCGCTCTTCTAGTGGCG 60
Db 1 AAGTGCTAAAGCTCCGAGGCCAAGCGCGCTGCTACTGCGCGCGCTCTTCTAGTGGCG 60
Qy 61 CGTTCCCGCCCTGGGTTGTCAACCGCGCGCGCGCGCGAGGAAGCACTGCAACAGGACCG 120
Db 61 CGTTCCCGCCCTGGGTTGTCAACCGCGCGCGCGCGCGAGGAAGCACTGCAACAGGACCG 120
Qy 121 GAGTGAGGCGCGCGAGCATGAAGCGCGCGCGCGCGCGCTCATAGCGCACTGCGGACCG 180
Db 121 GAGTGAGGCGCGCGAGCATGAAGCGCGCGCGCGCGCTCATAGCGCACTGCGGACCG 180
Qy 181 TCCGCGCGCGCGCGCGGGGGAAGAAAATGCAACATGCGCAGCAATGGAACACAGAACAG 240
Db 181 TCCGCGCGCGCGCGCGGGGGAAGAAAATGCAACATGCGCAGCAATGGAACACAGAACAG 240
Qy 241 CTGGGTGTTGAGATATTTGAAATCTGCGACTGTGAGGAGAAATTTGAATCACAGGATCGG 300
Db 241 CTGGGTGTTGAGATATTTGAAATCTGCGACTGTGAGGAGAAATTTGAATCACAGGATCGG 300
Qy 301 CCTAAATTTGGAGCTTTTATGTTGAGCGGTATTTCTGGAGTCACTTAAAAAGCTGCTT 360
Db 301 CCTAAATTTGGAGCTTTTATGTTGAGCGGTATTTCTGGAGTCACTTAAAAAGCTGCTT 360
Qy 361 GCGATACCAAGAAATATCATGGCTACATGAGTGGCTTAAGGCCACCATGATTTTCATGTTT 420
Db 361 GCGATACCAAGAAATATCATGGCTACATGAGTGGCTTAAGGCCACCATGATTTTCATGTTT 420
Qy 421 GTGAAGAGGAATGATCCAGATGGAATCTTTCAGACAGAACTCTATACCTTGGCATGCT 480
Db 421 GTGAAGAGGAATGATCCAGATGGAATCTTTCAGACAGAACTCTATACCTTGGCATGCT 480
Qy 481 GGTGAGAACAGAGAAATACCTGTTTATTTCTGAAATCTCCAAAATCTCAATAGAGCA 540
Db 481 GGTGAGAACAGAGAAATACCTGTTTATTTCTGAAATCTCCAAAATCTCAATAGAGCA 540
Qy 541 GCAGTCTTAAATGCTCTCTTTGGAAGCGCTCTTTTGGATCTTTTTCAGGCAACTGACTAT 600
Db 541 GCAGTCTTAAATGCTCTCTTTGGAAGCGCTCTTTTGGATCTTTTTCAGGCAACTGACTAT 600

QY	601	GGAAATGATTTCTCGAAGAGAGAACTATTAAAGAGAAAGAAACGCAATTTGAAACAGTCGGA	660	1681	TGTCCTATCAAGAGAGATAGCAATTAACAGTGGTGAATGGAGATTTCTTGGCGGCAT	1740
DB	601	GGAAATGATTTCTCGAAGAGAGAACTATTAAAGAGAAAGAAACGCAATTTGAAACAGTCGGA	660	1681	TGTCCTATCAAGAGAGATAGCAATTAACAGTGGTGAATGGAGATTTCTTGGCGGCAT	1740
QY	661	ATTGCTTCTTACGATTTATCACCAGAGAGTGAACATTTCTGTTTCAAGCCGCTAGTGA	720	1741	GGATCTAATATCCAAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGACCAACAAGAC	1800
DB	661	ATTGCTTCTTACGATTTATCACCAGAGAGTGAACATTTCTGTTTCAAGCCGCTAGTGA	720	1741	GGATCTAATATCCAAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGACCAACAAGAC	1800
QY	721	ATTATATCAGCTAAAGATGGAGGGCCACAAGATTTTACGCAACCACTTTTAAAGCCCAAT	780	1801	TCCTCTTTAGAGCATCACCTGTACGTACGTACGTAAATTCGTGAAGTGGAGAGGTGACAAGG	1860
DB	721	ATTATATCAGCTAAAGATGGAGGGCCACAAGATTTTACGCAACCACTTTTAAAGCCCAAT	780	1801	TCCTCTTTAGAGCATCACCTGTACGTACGTACGTAAATTCGTGAAGTGGAGAGGTGACAAGG	1860
QY	781	CTAGTGGAAACTAGTTGTCCCAACATACGGAATGGATCCAAATTTATGCCCCGCTGATCCA	840	1861	CTGACTACCGCTGCTACTCACATTTCTGCTGCATCAGTCAGCAGCTGTGACTTCTTTATA	1920
DB	781	CTAGTGGAAACTAGTTGTCCCAACATACGGAATGGATCCAAATTTATGCCCCGCTGATCCA	840	1861	CTGACTACCGCTGCTACTCACATTTCTGCTGCATCAGTCAGCAGCTGTGACTTCTTTATA	1920
QY	841	GACTGGATTCCTTTATACATAGCAACGATATTTGGATATCTTAACATCGTAAACAGAGAA	900	1921	AGTAAGTATAGTAAACAGAGAAATCCACACTGTGTGCTCCCTTTACAAGCTATCAAGTCTT	1980
DB	841	GACTGGATTCCTTTATACATAGCAACGATATTTGGATATCTTAACATCGTAAACAGAGAA	900	1921	AGTAAGTATAGTAAACAGAGAAATCCACACTGTGTGCTCCCTTTACAAGCTATCAAGTCTT	1980
QY	901	GAAAGGAGACTCACTTATGTGCAATGAGCTAGCAACATGGAAGAGATGCCAGATCA	960	1981	GAAGATGACCCAACTTTGCAAAAAGAAATTTTGGGCCACCAATTTTGGATTCAGCAGGT	2040
DB	901	GAAAGGAGACTCACTTATGTGCAATGAGCTAGCAACATGGAAGAGATGCCAGATCA	960	1981	GAAGATGACCCAACTTTGCAAAAAGAAATTTTGGGCCACCAATTTTGGATTCAGCAGGT	2040
QY	961	GCTGGAGTCCGCTACCTTTGTTCTCCAAAGAAATTTGATAGATATTTGGCTATTTGGTGG	1020	2041	CCTCTTCTGACTATCTCTCCAGAAATTTCTCTTTTGAAGTACTACTGGAATTTACA	2100
DB	961	GCTGGAGTCCGCTACCTTTGTTCTCCAAAGAAATTTGATAGATATTTGGCTATTTGGTGG	1020	2041	CCTCTTCTGACTATCTCTCCAGAAATTTCTCTTTTGAAGTACTACTGGAATTTACA	2100
QY	1021	TGTCAAAAGCTGAACAACCTCCAGTGTGGTAAATTTCTTAAGATTTCTATATGAAGAA	1080	2101	TTGTATCGGATGCTCTACAGGCTCATGATCTACAGCCTGGAAAGAAATATCTACTGTG	2160
DB	1021	TGTCAAAAGCTGAACAACCTCCAGTGTGGTAAATTTCTTAAGATTTCTATATGAAGAA	1080	2101	TTGTATCGGATGCTCTACAGGCTCATGATCTACAGCCTGGAAAGAAATATCTACTGTG	2160
QY	1081	AATGATGAATCTGAGGTGGAATTTATCATGTTATCATCCCTATGTTGGAAACAGGAGG	1140	2161	CTGTTTCATATATGTTGGTCTCTCAGGTGTCAGTTGGTGAATTAATTCGGTTTAAAGGAGTCAAG	2220
DB	1081	AATGATGAATCTGAGGTGGAATTTATCATGTTATCATCCCTATGTTGGAAACAGGAGG	1140	2161	CTGTTTCATATATGTTGGTCTCTCAGGTGTCAGTTGGTGAATTAATTCGGTTTAAAGGAGTCAAG	2220
QY	1141	GCAGATTCATTCGGTTATCTTAAACAGGTACAGCAAAATCTTAAAGTCACCTTTTAAAGATG	1200	2184	TAATTTCCGCTGGAATACCCCTAGCCTCTCTAGGTTATGTTGGTTAGTGTAGTAGACACAGG	2280
DB	1141	GCAGATTCATTCGGTTATCTTAAACAGGTACAGCAAAATCTTAAAGTCACCTTTTAAAGATG	1200	2184	TAATTTCCGCTGGAATACCCCTAGCCTCTCTAGGTTATGTTGGTTAGTGTAGTAGACACAGG	2280
QY	1201	TCAGAAATATGATTTGATGCTGAAGGAGGATCATAGATGTCATAGATAGAGAACTAAT	1260	2281	GGATCTGTCAACCGAGGCTTAAATTTGAAGGGCCCTTTAAATATAAAAAATGGGTCAAATA	2340
DB	1201	TCAGAAATATGATTTGATGCTGAAGGAGGATCATAGATGTCATAGATAGAGAACTAAT	1260	2281	GGATCTGTCAACCGAGGCTTAAATTTGAAGGGCCCTTTAAATATAAAAAATGGGTCAAATA	2340
QY	1261	CAACCTTTTGAAGTTCATTTTGAAGGAGTGAATATATTTGCCAGAGCTGGATGACTCCT	1320	2341	GAAATTCAGCATCAGGTGGAAGGATCCCAATATCTAGCTTTCTCGATATGATTTTCAATGAC	2400
DB	1261	CAACCTTTTGAAGTTCATTTTGAAGGAGTGAATATATTTGCCAGAGCTGGATGACTCCT	1320	2341	GAAATTCAGCATCAGGTGGAAGGATCCCAATATCTAGCTTTCTCGATATGATTTTCAATGAC	2400
QY	1321	GAGGAAATATGCTGGTCCATCTCTACTAGATCGCTCCAGACTCGCTTACAGATAGTG	1380	2401	TTAGATCGTGTGGGCATCCACGGCTGCTCTATGGAGGATACCTCTCCCTGATGGCATTA	2460
DB	1321	GAGGAAATATGCTGGTCCATCTCTACTAGATCGCTCCAGACTCGCTTACAGATAGTG	1380	2401	TTAGATCGTGTGGGCATCCACGGCTGCTCTATGGAGGATACCTCTCCCTGATGGCATTA	2460
QY	1381	TTGATCTCACCTGAATTTATTTATCCAGTAGAAGATGATGTTATGAAAGGACAGAGACTC	1440	2461	ATGCAGAGGTCAGATATCTTTCAGGGTTCCTATTTGCTGGGGCCCGCCAGTCACTCTGTGGATC	2520
DB	1381	TTGATCTCACCTGAATTTATTTATCCAGTAGAAGATGATGTTATGAAAGGACAGAGACTC	1440	2461	ATGCAGAGGTCAGATATCTTTCAGGGTTCCTATTTGCTGGGGCCCGCCAGTCACTCTGTGGATC	2520
QY	1441	ATTGAGTCAGTGCCTGATTTCTGTGACGCACTAAATTTATCTATGAAGAAACACAGACATC	1500	2521	TTCTATGATACAGGATACCGGAACTGTTATATGGGTTCACCTGACAGAAATGAACAGGGC	2580
DB	1441	ATTGAGTCAGTGCCTGATTTCTGTGACGCACTAAATTTATCTATGAAGAAACACAGACATC	1500	2521	TTCTATGATACAGGATACCGGAACTGTTATATGGGTTCACCTGACAGAAATGAACAGGGC	2580
QY	1501	TGGATAATATCCATGACATCTTTTCATGTTTCCCAAGTCAGGAAGGAAATTTGAG	1560	2281	TATTACTTAGGATCTGTGGCCATGCAAGCAGAAAGTTCCCTCTCTGAAACCAATCGTTTA	2340
DB	1501	TGGATAATATCCATGACATCTTTTCATGTTTCCCAAGTCAGGAAGGAAATTTGAG	1560	2281	TATTACTTAGGATCTGTGGCCATGCAAGCAGAAAGTTCCCTCTCTGAAACCAATCGTTTA	2340
QY	1561	TTTATTTTGGCTCTGAATGCAAAAACAGGTTTCCGTCATTTATATACAAAATTTACATCTATT	1620	2641	CTGCTCTTACATGTTTCTTGGATGAGATGTCATTTTGGCATATACCAAGTATATTAATG	2700
DB	1561	TTTATTTTGGCTCTGAATGCAAAAACAGGTTTCCGTCATTTATATACAAAATTTACATCTATT	1620	2641	CTGCTCTTACATGTTTCTTGGATGAGATGTCATTTTGGCATATACCAAGTATATTAATG	2700
QY	1621	TTAAAGGAAAGCAATATATAACGATCCAGTGGTGGCTGCTGCTCCAAAGTGAATTCAG	1680	2701	AGTTTCTTCTGAGGGCTGGAAAGCCATATGATTTTACAGATCTATCTCTCAGGAGAGACAC	2760
DB	1621	TTAAAGGAAAGCAATATATAACGATCCAGTGGTGGCTGCTGCTCCAAAGTGAATTCAG	1680	2701	AGTTTCTTCTGAGGGCTGGAAAGCCATATGATTTTACAGATCTATCTCTCAGGAGAGACAC	2760
QY	1621	TTAAAGGAAAGCAATATATAACGATCCAGTGGTGGCTGCTGCTCCAAAGTGAATTCAG	1680	2761	AGCATAGAGTTCCTGAAATCGGAGAAACATTTATGAACTGCTATCTTTTGGCATCTCACTCA	2820

Db	2461	AGCATAAGAGTTCCTGAATCGGAGAACATTATGAACCTGCTCTTTGCACTACCTTCAA	2520
Qy	2821	GAATAACCTTGGATCACGCTATTGCTGCTCTAAAAAGTGATATAATTTTGACCTGTGTAGAAC	2880
Db	2521	GAATAACCTTGGATCACGCTATTGCTGCTCTAAAAAGTGATATAATTTTGACCTGTGTAGAAC	2580
Qy	2881	TCTCTGGTATACACTGGCTATTTAAACCAATGAGGAGGTTTAAATCAAACAGAAAAACACAGA	2940
Db	2581	TCTCTGGTATACACTGGCTATTTAAACCAATGAGGAGGTTTAAATCAAACAGAAAAACACAGA	2640
Qy	2941	ATTGATCATCATTTTGGATACCTGCGCATGTAAACATCTACTCCTGAAAAATAAATGTGGTG	3000
Db	2641	ATTGATCATCATTTTGGATACCTGCGCATGTAAACATCTACTCCTGAAAAATAAATGTGGTG	2700
Qy	3001	CCATGCGGGGTCTACGGTTTGTGGTAGTAATCTAATACCTTAACCCACATGCTCAAAA	3060
Db	2701	CCATGCGGGGTCTACGGTTTGTGGTAGTAATCTAATACCTTAACCCACATGCTCAAAA	2760
Qy	3061	TCAATGTACATATTTCTGTGAGAGACCCGCGCATACCATAAGAAATTACTAAAAA	3120
Db	2761	TCAATGTACATATTTCTGTGAGAGACCCGCGCATACCATAAGAAATTACTAAAAA	2820

Search completed: April 15, 2006, 03:22:29
Job time : 2427 secs

Result No.	Score	Query		DB	ID	Description
		Match	%			
1	3105.2	99.5	3143	14	US-11-151-601-19	Sequence 19, Appl
2	2645.8	84.8	2649	14	US-11-151-601-21	Sequence 21, Appl
C 3	1118	3.8	1346	6	US-09-925-065A-669313	Sequence 669313, Appl
C 4	113.6	3.6	609	6	US-09-925-065A-743558	Sequence 743558, Appl
C 5	113.6	3.6	609	6	US-09-925-065A-743559	Sequence 743559, Appl
						Sequence 1186, Ap
6	79.4	3.0	2238	11	US-11-079-463-1186	Sequence 11, Appl
7	79.4	2.5	3332	11	US-11-208-288-1	Sequence 34, Appl
8	79.4	2.5	3407	9	US-10-501-035-34	Sequence 3, Appl
9	77.8	2.5	2217	11	US-11-208-288-3	Sequence 1, Appl
10	77.8	2.5	2301	8	US-10-523-789-1	Sequence 1, Appl
C 11	73.4	2.4	612	6	US-09-925-065A-818446	Sequence 818446, Appl
12	71.6	2.3	4852	14	US-11-136-527-2130	Sequence 2130, Ap
13	71.1	2.3	2283	11	US-11-208-288-5	Sequence 5, Appl
14	68.6	2.2	1884	11	US-11-079-463-2536	Sequence 2536, Ap
15	65.8	2.1	2814	11	US-11-245-147-168	Sequence 168, App
16	65.8	2.1	2814	14	US-11-186-284-54	Sequence 54, Appl
17	63.2	2.0	535	11	US-11-226-869-428	Sequence 428, App
18	52.4	1.7	2778	9	US-10-933-182A-5649	Sequence 5649, Ap

		LENGTH: 3143	
		TYPE: DNA	
		ORGANISM: Homo sapiens	
		FEATURE:	
		NAME/KEY: CDS	
		LOCATION: (229)...(2874)	
		US-11-151-601-19	
		Query Match	99.5%; Score 3105.2; DB 14; Length 3143;
		Best Local Similarity	99.9%; Pred. No. 0;
		Matches 3107; Conservative	0; Mismatches 3; Indels 0; Gaps 0;
Qy	11	GCCTCCGAGGCCAAGCCGCTGCTACTGCGCGCGCTCTTCTTAGTGCGCGCTGCGCGC	70
Db	26	GCCTCCGAGGCCAAGCCGCTGCTACTGCGCGCGCTCTTCTTAGTGCGCGCTGCGCGC	85
Qy	71	CTGGGTGTGCACCGCGCGCGCGCGAGGAAGCCACTGCAACGAGGACCGGAGTGAGGC	130
Db	86	CTGGGTGTGCACCGCGCGCGCGCGAGGAAGCCACTFACAACGAGGACCGGAGTGAGGC	145
Qy	131	GGCGCAGCATGAAGCGCGCAGGCGCGCTCCATAGCGCAGTCCGGACGCTCGGCGCGG	190
Db	146	GGCGCAGCATGAAGCGCGCAGGCGCGCTCCATAGCGCAGTCCGGACGCTCGGCGCGG	205
Qy	191	GCGGGGGGAAGGAATGCAACATGGCAGCAGCAATGGAACAGAACAGCTGGGTGTG	250
Db	206	GCGGGGGGAAGGAATGCAACATGGCAGCAGCAATGGAACAGAACAGCTGGGTGTG	265
Qy	251	AGATATTTGAACTCGGACCTGTGAGGAGAAATATTGAATCAGAGATCGGCCCTAAATTGG	310
Db	266	AGATATTTGAACTCGGACCTGTGAGGAGAAATATTGAATCAGAGATCGGCCCTAAATTGG	325
Qy	311	AGCCCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGAGCTGCTTGGCGATPACA	370
Db	326	AGCCCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGAGCTGCTTGGCGATPACA	385
Qy	371	GAAAAATATCATGGCTACATGATGCTAAGGCACCAATGATTTTCATGTTTGTGAAGAGGA	430
Db	386	GAAAAATATCATGGCTACATGATGCTAAGGCACCAATGATTTTCATGTTTGTGAAGAGGA	445
Qy	431	ATGATCCAGATGGAACCTCATTCAGACAGAACTCTATTACCTTGGCCATGCTCTGGTGAGAA	490
Db	446	ATGATCCAGATGGAACCTCATTCAGACAGAACTCTATTACCTTGGCCATGCTCTGGTGAGAA	505
Qy	491	GAGAAATACACTGTTTATTCTGAAATTTCCCAAACTATCAATAGAGCAGCAGTCTTAA	550
Db	506	GAGAAATACACTGTTTATTCTGAAATTTCCCAAACTATCAATAGAGCAGCAGTCTTAA	565
Qy	551	TGCTCTCTTGGAGGCTCTTTTGGATCTTTTTCAGGCAACACTGGACTATGGAATGTATT	610
Db	566	TGCTCTCTTGGAGGCTCTTTTGGATCTTTTTCAGGCAACACTGGACTATGGAATGTATT	625
Qy	611	CTCGAGAGAGAACTATTAGAGAAAGAAACGATTTGGAACAGTCGGAAATGCTTCTT	670
Db	626	CTCGAGAGAGAACTATTAGAGAAAGAAACGATTTGGAACAGTCGGAAATGCTTCTT	685
Qy	671	ACGATTAACAAGAAAGTGGAACATTTCTTGTTTCAAGCGGTAGTGGAAATTTATCAGC	730
Db	686	ACGATTTATCACCAAGAAAGTGGAACATTTCTTGTTTCAAGCGGTAGTGGAAATTTATCAGC	745
Qy	731	TAAAGATGAGGGCCACAAGATTTACGCAACAACTTTAAGCCCAATCTAGTGGA	790
Db	746	TAAAGATGAGGGCCACAAGATTTTACGCAACAACTTTAAGCCCAATCTAGTGGA	805
Qy	791	CTAGTTGTCCCAACATACGGATGGATCCAAATTTATGCCCGCTGATCCAGACTGGATTG	850
Db	806	CTAGTTGTCCCAACATACGGATGGATCCAAATTTATGCCCGCTGATCCAGACTGGATTG	865
Qy	851	CTTTTATACATAGCAACATATTTGGATATCTAACATCGTAAACAGAGAAAGAGAC	910
Db	866	CTTTTATACATAGCAACATATTTGGATATCTAACATCGTAAACAGAGAAAGAGAC	925
Qy	911	TCACCTTATGTGCAATGAGCTAGCCAACTATGGAAGAAAGATGCCAGATCAGCTGGAGTCG	970

Db	926	TCACCTTATGTGCAATGAGCTAGCCAACTAGGAAAGATGCCAGATCAGCTGGAGTCG	985
Qy	971	CTACCTTTGTTCTCCAGAGAAATTTGATAGATATCTGGCTATTTGGTGTGTCCTCAAG	1030
Db	986	CTACCTTTGTTCTCCAGAGAAATTTGATAGATATCTGGCTATTTGGTGTGTCCTCAAG	1045
Qy	1031	CTGAAACAACTCCGAGTGGTAAATCTTAGAATCTATATGAAGAAATATGATGAAT	1090
Db	1046	CTGAAACAACTCCGAGTGGTAAATCTTAGAATCTATATGAAGAAATATGATGAAT	1105
Qy	1091	CTGAGGTGAAATTTATCATGTTACATCCCTATTTGGGAAACAAAGAGGCGCAGATTCAT	1150
Db	1106	CTGAGGTGAAATTTATCATGTTACATCCCTATTTGGGAAACAAAGAGGCGCAGATTCAT	1165
Qy	1151	TCCGTTATCTCTAAAAACAGGTACAGCAATCCCTAAAGTCACCTTTTAAAGATGTCAGAAATA	1210
Db	1166	TCCGTTATCTCTAAAAACAGGTACAGCAATCCCTAAAGTCACCTTTTAAAGATGTCAGAAATA	1225
Qy	1211	TGATTGATCTCAAGAGGAGGATCATAGATGTATAGATTAAGGAACTAAATTCAACTTTTG	1270
Db	1226	TGATTGATCTCAAGAGGAGGATCATAGATGTATAGATTAAGGAACTAAATTCAACTTTTG	1285
Qy	1271	AGATTCTATTTGAAGGAGTTGAATATATTGCCAGAGCTGGATGAGCTCCTGAGGGAAT	1330
Db	1286	AGATTCTATTTGAAGGAGTTGAATATATTGCCAGAGCTGGATGAGCTCCTGAGGGAAT	1345
Qy	1331	ATGCTTGTGCTCATCTAGATCGCTCCAGACTCCGCTACAGATAGTGTGATCTCAC	1390
Db	1346	ATGCTTGTGCTCATCTAGATCGCTCCAGACTCCGCTACAGATAGTGTGATCTCAC	1405
Qy	1391	CTGATTTATTTATCCAGTAGAAGATGATGTTATGGAAGGACAGAGACTCAATTTGAGTCAG	1450
Db	1406	CTGATTTATTTATCCAGTAGAAGATGATGTTATGGAAGGACAGAGACTCAATTTGAGTCAG	1465
Qy	1451	TGCTCTGATTTCTGAGCGCCATTAATCTATGAAGAAACAAACAGACATCTGGATAAATA	1510
Db	1466	TGCTCTGATTTCTGAGCGCCATTAATCTATGAAGAAACAAACAGACATCTGGATAAATA	1525
Qy	1511	TCCATGACATCTTTTCAATGTTTTCCTCAAGTCAGAAAGAGAAATTTGAGTTATTTTG	1570
Db	1526	TCCATGACATCTTTTCAATGTTTTCCTCAAGTCAGAAAGAGAAATTTGAGTTATTTTG	1585
Qy	1571	CTCTCTGAAATGCAAAACAGGTTTCCGTCAATTTATACAAATTTACATCTATTTTAAAGAAA	1630
Db	1586	CTCTCTGAAATGCAAAACAGGTTTCCGTCAATTTATACAAATTTACATCTATTTTAAAGAAA	1645
Qy	1631	GCAATATATAACGATCCAGTGGTGGCTGCTGCTCCAAAGTGATTTCAAAGTGTCTATCA	1690
Db	1646	GCAATATATAACGATCCAGTGGTGGCTGCTGCTCCAAAGTGATTTCAAAGTGTCTATCA	1705
Qy	1691	AAGAGGAGATAGCAATTAACAGTGGTGAATGGGAAGTTCTTTGGCGGCATGATCTAATA	1750
Db	1706	AAGAGGAGATAGCAATTAACAGTGGTGAATGGGAAGTTCTTTGGCGGCATGATCTAATA	1765
Qy	1751	TCCAAAGTTCAGTACGAGAGGCTGGTATATTTTGAAGGCACCAAGAGCTCCCTTTAG	1810
Db	1766	TCCAAAGTTCAGTACGAGAGGCTGGTATATTTTGAAGGCACCAAGAGCTCCCTTTAG	1825
Qy	1811	AGCATCACTGTACGTAGTCAGTTACGTAAATCTCGGAGAGCTGACAGGCTGACTGACC	1870
Db	1826	AGCATCACTGTACGTAGTCAGTTACGTAAATCTCGGAGAGCTGACAGGCTGACTGACC	1885
Qy	1871	GTGGCTATCTACATCTTCTGCTGCATCAGTCAGACTGTGACTTCTTTTAAAGTAAGTATA	1930
Db	1886	GTGGCTATCTACATCTTCTGCTGCATCAGTCAGACTGTGACTTCTTTTAAAGTAAGTATA	1945
Qy	1931	GTAAACAGAGAAATCCACACTGTGTGCTCCCTTTTCAAGCTATCAAGTCTCTGAAGATGACC	1990
Db	1946	GTAAACAGAGAAATCCACACTGTGTGCTCCCTTTTCAAGCTATCAAGTCTCTGAAGATGACC	2005
Qy	1991	CAACTTGGCAAAACAAAGGAATTTTGGGCGCACCATTTTCGATTTACGACAGGTCCTCTCTG	2050

Db 2006 CAACCTTGCAAAACAAAGGAATTTTGGGCCACCAATTTTGGATTGAGGAGTCTCTCTCTG 2065
Qy 2051 ACTATACCTCTCCAGAAATTTTCTTTTGAAGTACTACTGATTTACATTTGATGGA 2110
Db 2066 ACTATACCTCTCCAGAAATTTTCTTTTGAAGTACTACTGATTTACATTTGATGGA 2125
Qy 2111 TGCTCTACAAGCCTCATGATCTACAGCCTGGAAGAAATATCTACTGTGCTTCAAT 2170
Db 2126 TGCTCTACAAGCCTCATGATCTACAGCCTGGAAGAAATATCTACTGTGCTTCAAT 2185
Qy 2171 ATGGTGGTCTCAGGTGAGTGTGTAATATGTTTAAAGAGTCAAGTATTTCCGCT 2230
Db 2186 ATGGTGGTCTCAGGTGAGTGTGTAATATGTTTAAAGAGTCAAGTATTTCCGCT 2245
Qy 2231 TGAATACCTGACCTCTAGTGTATGTTGTAGTGATAGACACAGGATCTGTG 2290
Db 2246 TGAATACCTGACCTCTAGTGTATGTTGTAGTGATAGACACAGGATCTGTG 2305
Qy 2291 ACCGAGGGCTTAAATTTGAAGCGCTTTTAAATATAAATGGTCAAAATAGAAATGAG 2350
Db 2306 ACCGAGGGCTTAAATTTGAAGCGCTTTTAAATATAAATGGTCAAAATAGAAATGAG 2365
Qy 2351 ATCAGGTGGAAGGACTCCAATATCTAGCTTCTGATGATATTTCAATGATCTAGATG 2410
Db 2366 ATCAGGTGGAAGGACTCCAATATCTAGCTTCTGATGATATTTCAATGATCTAGATG 2425
Qy 2411 TGGGATCCAGGCTGCTTATGAGGATACCTCTCCCTGATGATTAATGAGGATTAATGAG 2470
Db 2426 TGGGATCCAGGCTGCTTATGAGGATACCTCTCCCTGATGATTAATGAGGATTAATGAG 2485
Qy 2471 CAGATATCTTCAAGGCTGCTTATGAGGATACCTCTCCCTGATGATTAATGAGGATTAATGAG 2530
Db 2486 CAGATATCTTCAAGGCTGCTTATGAGGATACCTCTCCCTGATGATTAATGAGGATTAATGAG 2545
Qy 2531 CAGGATACCGGAAGCTTATATGAGGATACCTCTCCCTGATGATTAATGAGGATTAATGAG 2590
Db 2546 CAGGATACCGGAAGCTTATATGAGGATACCTCTCCCTGATGATTAATGAGGATTAATGAG 2605
Qy 2591 GATCTGTGGCCATGCAAGGAAATTTCCCTCTGAAACCAATCGTTTACTCTCTTAC 2650
Db 2606 GATCTGTGGCCATGCAAGGAAATTTCCCTCTGAAACCAATCGTTTACTCTCTTAC 2665
Qy 2651 ATGGTTCTGATGAGATGCTCAATTTTGACATACAGTATATTAATGATTTTCTGAT 2710
Db 2666 ATGGTTCTGATGAGATGCTCAATTTTGACATACAGTATATTAATGATTTTCTGAT 2725
Qy 2711 TGAGGCTGGAAGCCATATGATTTACAGATCTATCTCTCAGGAGAGACACAGATTAAG 2770
Db 2726 TGAGGCTGGAAGCCATATGATTTACAGATCTATCTCTCAGGAGAGACACAGATTAAG 2785
Qy 2771 TTCTGAAATCGGAGAAACATTAATGAACTGATCTTTTGCATCTACCTTCAAGAAACCTTG 2830
Db 2786 TTCTGAAATCGGAGAAACATTAATGAACTGATCTTTTGCATCTACCTTCAAGAAACCTTG 2845
Qy 2831 GATCAGTATTTCTGCTTAAAGGATATATTTTGACCTGCTGATGAACTCTCTGAT 2890
Db 2846 GATCAGTATTTCTGCTTAAAGGATATATTTTGACCTGCTGATGAACTCTCTGAT 2905
Qy 2891 ACATGCTTATTTAAACCAATGAGGAGTTTTAAATCAACAGAAACACAGAAATGATCATC 2950
Db 2906 ACATGCTTATTTAAACCAATGAGGAGTTTTAAATCAACAGAAACACAGAAATGATCATC 2965
Qy 2951 ACATTTTGTATCTGCTGATTAACATCTATCTCTGAAATTAATGTTGGTGCATGAGGG 3010
Db 2966 ACATTTTGTATCTGCTGATTAACATCTATCTCTGAAATTAATGTTGGTGCATGAGGG 3025
Qy 3011 GTCTACGTTTGTGTTAGTAAATCTAATACCTTTAAACCCACATGCTCAAAATCAATGATA 3070
Db 3026 GTCTACGTTTGTGTTAGTAAATCTAATACCTTTAAACCCACATGCTCAAAATCAATGATA 3085
Qy 3071 CATATTTCTGAGAGCCAGCAATACCAATGATTTACTTAAAAA 3120
Db 3086 CATATTTCTGAGAGCCAGCAATACCAATGATTTACTTAAAAA 3135

RESULT 2
US-11-151-601-21
; Sequence 21, Application US/11151601
; Publication No. US20060003413A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Olandt, Peter J.
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Williamson, Mark
; APPLICANT: Weich, Nadine
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASE, PHOSPHATASE,
; AND PROTEASE FAMILY MEMBERS AND USES THEREOF
; FILE REFERENCE: MPI00-054PIRCP10MNDIVM
; CURRENT APPLICATION NUMBER: US/11/151,601
; CURRENT FILING DATE: 2005-06-13
; PRIOR APPLICATION NUMBER: US 10/170,789
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: US 09/797,039
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: PCT/US01/06525
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/186,061
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 09/882,166
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: PCT/US01/19269
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/212,078
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/934,406
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: PCT/US01/26052
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/226,740
; PRIOR FILING DATE: 2000-08-21
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 2649
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-151-601-21

Query Match 84.8%; Score 2645.8; DB 14; Length 2649;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2647; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 214 ATGCGACGACCAATGAAACAGACAGCAGTGGGTGGAGATATTTGAACTGCGACCTGT 273
Db 1 ATGCGACGACCAATGAAACAGACAGCAGTGGGTGGAGATATTTGAACTGCGACCTGT 60
Qy 274 GAGGAGAAATTTCAATCAGAGATCGCTTAAATTTGAGCCCTTTTATGTTGAGCGTAT 333
Db 61 GAGGAGAAATTTCAATCAGAGATCGCTTAAATTTGAGCCCTTTTATGTTGAGCGTAT 120
Qy 334 TCCTGGAGTCAGCTTAAAGCTGCTTGGCAGATCCAGAAATATCATGGCTACATGATG 393
Db 121 TCCTGGAGTCAGCTTAAAGCTGCTTGGCAGATCCAGAAATATCATGGCTACATGATG 180
Qy 394 GCTAAGGACCAATGATGATTTTCAATGTTGAGAGGATGATCCAGATGGACCTCATCA 453
Db 181 GCTAAGGACCAATGATGATTTTCAATGTTGAGAGGATGATCCAGATGGACCTCATCA 240
Qy 454 GACAGATCTTATACCTTGCATGCTGTTGAGAAACAGAGAAATATCACTGTTTATTTCT 513
Db 241 GACAGATCTTATACCTTGCATGCTGTTGAGAAACAGAGAAATATCACTGTTTATTTCT 300
Qy 514 GAAATTTCCCAAACTATCAATAGACAGCAGTCTTAAATGCTCTCTTTGGAGCCTCTTTTG 573


```
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTERIOIDES FR
; FILE REFERENCE: PATH00-03DIV2
; CURRENT APPLICATION NUMBER: US/11/079,463
; CURRENT FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/128,705
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: US 09/540,209
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 1186
; LENGTH: 2238
; TYPE: DNA
; ORGANISM: B. fragilis
; US-11-079-463-1186

Query Match      3.0%; Score 94.8; DB 11; Length 2238;
Best Local Similarity 48.4%; Pred. No. 5.1e-15;
Matches 369; Conservative 0; Mismatches 372; Indels 21; Gaps 3;

Qy 2064 AGAAATTTCTCTTTGAAAGTACTACTGGATTTTACATTTGATGGGATGCTCTACAGCC 2123
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1491 AGAAATTTCTCTTTGAAAGTACTACTGGATTTTACATTTGATGGGATGCTCTACAGCC 1550
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2124 TCATGATCTACAGCTGGAAGAAATATCTTACTGTCTGTTATATATGTTGGTCTCTCA 2183
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1551 GGTCAATTCGATCTGCGCAACGTTATCCGTTACTGATGTTCCAGTATAGCGGTCGGG 1610
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2184 GGTGCAATTCGATCTGCGTTTAAAGGAGTCAAGTATTTCCGTTTGAATACCTAGC 2243
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1611 TTCGCAACAGGTTCTGGCAAAATGGGAATCAGTTGGGAACCTACATG-----GC 1661
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2244 CTCTAGCTTATGTTGTTAGTATAGACACACAGGGATCTCTGACCGAGGCTTAA 2303
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1662 GAGCCTCGGTTACGTTGTTAGTCTGTTGATAGTGGTGGGACAGGTGGCGTGGCAGTA 1721
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2304 ATTTGAAGCGGCTTTAAATATAAATGSGTCAAAATGAATTAACGATCAGGTGGAAGG 2363
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1722 ATTCCAGAAATGCACCTACCTGAACCTGGGTGTAAGAGAGCTAAAGACAGGTGGAAGC 1781
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2364 ACTCCAAATATCTAGCTTCTCGATATGATTTCAITGACTTATAGTCTGTTGGGCATCCAG 2423
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1782 TGCCAAATATCTGGGTGGACTGCTTATGTG---GACAAAGGACGTATTTGGTATCTGGGG 1838
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2424 CTGGTCTATGAGGATACCTCTCCCTGATGGCATTAAATGACAGAGTCAAGATCTTCAG 2483
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1839 ATGGAGTTTCGGCGGATATATGACCATCATGATGATGAGCGAAGGTACACCCGTGTTAA 1898
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2484 GGTGCTATTGCTGGGGCCCACTCACTCTGTGGATCTTCTATGATACAGGATACACGGA 2543
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1899 AGCGGAGTTGCTGTGGCGCACCTACAGACTGGAATATTTACGATACAGTATATACCGA 1958
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2544 ACCTTATATGGGTCACTCCCTGACCAAGTGAACAGGCTTATTTACTTAGGATCTGTGGCCAT 2603
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1959 ACCTTTATGCGCACCGCGAAGAAATGCCGAGGCTATAAAGCAGCTTCAGATTCAG 2018
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2604 GCAACGAGAAAGTCCCTCTGACCAATCGTTTACTGCTCTTACATGTTTCTGGA 2663
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2019 CCGTCAGACACCTGCTGATGGTACC-----TGCTCTTGTACCGTATGSCAG 2069
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2664 TGAGATGTCATTTTGGCACATACAGTATATTTACTGATGTTTTTATGAGGCTCGAAA 2723
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2070 TGATAATGTTCACTTCCAGAACTGTACAGAAATATGACAGACACCTGGTAACTCCGAAA 2129
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2724 GCCATATGTTTACAGATCTATCTCTCAGGAGAGACACAGCATAAAGAGTTCTGAAATCGGG 2783
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2130 ACAGTTTCGATATGACGGTATACACCAACCGGAATCATAGCATCTATGTTGGAATACCCG 2189
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2784 AGAACTATGAACTGCATCTTTTTCACCTACTCTTCAAGAAA 2825
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2190 TAACCACTTGTATACGAAGCTGACGAACCTTCTTCGGAATAA 2231
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
RESULT 8
US-10-501-035-34
; Sequence 34, Application US/10501035
```

```
RESULT 7
US-11-208-288-1
; Sequence 1, Application US/11208288
; Publication No. US20060051366A1
; GENERAL INFORMATION:
; APPLICANT: CHANG, Chiwen
; TITLE OF INVENTION: USE OF SOLUBLE CD26 AS INHIBITOR OF
; FILE REFERENCE: 39533-0001
; TITLE OF INVENTION: ANGIOGENESIS AND INFLAMMATION
; CURRENT FILING DATE: 2005-08-18
; PRIOR APPLICATION NUMBER: US 60/605,013
; PRIOR FILING DATE: 2004-08-26
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3332
; TYPE: DNA
; ORGANISM: homo sapiens
; US-11-208-288-1
```

```
Query Match      2.5%; Score 79.4; DB 11; Length 3332;
Best Local Similarity 47.7%; Pred. No. 1.2e-10;
Matches 296; Conservative 0; Mismatches 321; Indels 4; Gaps 2;

Qy 2270 TAGACAACACAGGGGATCCTGTCCACCGAGGGCTTTAAATTTGAAGGCGCCCTTTAAATATATAA 2329
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1733 TTGATGCGCAGAGGAGTGGTTTACCAAGGAGATTAAGATCATGATGCAATCAACAGAGAC 1792
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2330 TGGGTCAAATAGAAATTTGACGATCAGGTGGAGGACTCCAATATCTAGCTTCTCGATATG 2389
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1793 TGGGAACATTTGAAGTTGAAGATCAATTTGAAGCAGCCAGACAATTTTCAAAAATGG 1849
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2390 ATTTCATTCATTTAGATCTGTGGGCATCCACGGCTGCTTATGGAGGATACCTCTCCC 2449
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1950 GATTTGTGACAAACAAAGAAATTCGAATTTGGGCTGTCATATGGAGGTACGTAACT 1909
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2450 TGATGCAATTAATCAGAGTCAAGATATCTTCAGGTTGCTTATTTGCTGGGCGCCAGTCA 2509
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1910 CAATGCTCTGGGATCGGAGTGGGCTGTTCAAGTGGATAGCCGTGGGCGCTGTAT 1969
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2510 CTCTGTGATCTTCTATATACAGGATACAGGAACGTTTATGGTCACTCCCTGACACAGA 2569
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1970 CCGGTTGGGAGTACTATGACTCAGTGTACACAGAACGTTACATGGGTCTCCAACTCAG 2029
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2570 ATGAACAGGGCTATTATCTTAGGATCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGAAC 2629
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2030 AAGACAACCTTGACCAATTACAGAAATTCACAGTCAAGCAGAGCTGAAAATTTTAAAC 2089
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2630 CAAATCGTTTACTGCTCTTACATGTTTCTCGATGAGAAATGTCATTTTGCACATACCA 2689
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2090 AAGTTGAGTACCTCTCTTATTCATGGAACAGCAGATGATAACGTTCACTTTTCAGCAGTCAG 2149
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2690 GTATATTTACTGAGTTTTTTTAGTGAGGCTGGGAGGCTGAAATGTCATTTTGCATCTCTC 2749
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2150 CTCAGATCTCCAAAGCCCTGGTCTGATGTTGGAGTGGATTTCCAGGCAATGTGGTATAC 2209
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2750 AGGAGAGACACAGCATAAGAGTTCTCGAATCGGAGAACATTTATGACATGCTATTTTGC 2809
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2210 ATGAACACCATGGAATAGCTAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2269
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2810 ACTACCTCTCAGAAACCTTGGATCAGCTATGCTCTTAAAGTGATATAAATTTTGAAC 2869
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2270 ACTTCATAAACCAATGTTTCTCTTTACCT-TAGCACTTCAAAAATACCATGCAATTTAAAG 2328
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2870 CTGTGTAGAACTCTCTGGTAT 2890
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2329 CTTATTAAAAAATCAATTTTGT 2349
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Publication No. US20060046249A1
GENERAL INFORMATION:
APPLICANT: Bristol-Myers Squibb Company
TITLE OF INVENTION: IDENTIFICATION OF POLYNUCLEOTIDES AND POLYPEPTIDE FOR PREDICTING
FILE REFERENCE: 39533-0001
CURRENT APPLICATION NUMBER: US/11/208,288
CURRENT FILING DATE: 2005-08-18
PRIOR APPLICATION NUMBER: US 60/605,013
PRIOR FILING DATE: 2004-08-26
NUMBER OF SEQ ID NOS: 8
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 2217
TYPE: DNA
ORGANISM: Homo sapiens
US-10-501-035-34

Query Match 2.5%; Score 79.4; DB 9; Length 3407;
Best Local Similarity 47.7%; Pred. No. 1.3e-10;
Matches 296; Conservative 0; Mismatches 321; Indels 4; Gaps 2;

QY 2270 TAGACAAACAGGGGATCCTGTCAACGAGGCTTAAATTTGAAGGCGCTTTAAATATATAAA 2329
DB 1808 TTGATGCGCAGAGGAGTGGTTACCAAGGAGATAAGATCATGTCATCAATCAACAGAGAC 1867
QY 2330 TGGGTCAATAGAAATTTGACGATCAGGTGGAGGACTCCATATCTAGCTTCTCGATATG 2389
DB 1868 TGGGAACATTTGAAGTTGAAGATCAAAATTTGAAGGAGCCAGACAATTTTCAAAAATGG 1924
QY 2390 ATTTCATTGACTTGTAGATCTGTGGGATCCAGCGCTGTCTATGGAGGATACCTCTCCC 2449
DB 1925 GATTTGTGACAAACAAATTTGCAATTTGGGGCTGTCTATGGAGGATACGTAACCT 1984
QY 2450 TGATGCGATTAAATGACAGAGTTCAGATATCTTCAGGGTTGCTTATGCTGGGGCCCCAGTCA 2509
DB 1985 CAATGGTCTGGGATCGGGAAGTGGCGTGTCTCAAGTGTGGAATAGCGGTGGCGCTGTAT 2044
QY 2510 CTCTGTGGATCTTCTATATAGATACAGGATACAGGAAGCTTATATGGGTCAACCTGACAC 2569
DB 2045 CCGGTGGGAGTACTATGACTCAGTGTACACAGAACGTTATCATGGGTCTCCCAACTCCAG 2104
QY 2570 ATGACAGGGCTATTACTTAGGATCTGTGGCCATCGAGCAGAGAAAGTTCCCTCTGAAAC 2629
DB 2105 AAGACACCTTGACCATTAACAGAAATTCACAGTATGACAGAGCTGAAATTTTAAAC 2164
QY 2630 CAAATCGTTTACTGCTTACATGTTTCTCGATGAGAAATGTCCTATTTTGCACATACCA 2689
DB 2165 AAGTTGAGTACCTCTTATTCATGGAACAGCAGATGATACGTTTCACTTTTCAGCAGTCAG 2224
QY 2690 GTATATPACGTATTTTATGAGGGCTGGAAAGCCATATGATTTACAGATCTATCTCTC 2749
DB 2225 CTCAGATCTCAAAGCCCTGGTGCATGTTGGAGTGGATTTCCAGGCAATGTGGTATATCTG 2284
QY 2750 AGGAGACACACAGATAAGATCTCTGATCGGAGACATTAAGACTGCATCTTTTGC 2809
DB 2285 ATGAAGACCATGGAATAGCTAGCAGCAGCAGCAACATATATATACCCACATGAGCC 2344
QY 2810 ACTACCTTCAAGAAACCTTGGATCAGTATGCTGCTCTAAAGTGATATAAATTTGAC 2869
DB 2345 ACTTCATAAACAATGTTCTCTTTACCT-TAGCACCTCAAAATACCATGCTCAATTTAAAG 2403
QY 2870 CTGTGTAGAACTCTCTGGTAT 2890
DB 2404 CTTATTAAACTCATTTTGT 2424

RESULT 9
US-11-208-288-3
Sequence 3, Application US/11208288
Publication No. US20060051366A1
GENERAL INFORMATION:
APPLICANT: TANIABE SEIYAKU CO., LTD.
TITLE OF INVENTION: three-dimensional structure of dipeptidyl peptidase IV
FILE REFERENCE: 03-039-PCT
CURRENT APPLICATION NUMBER: US/10/522,789
CURRENT FILING DATE: 2005-01-28
PRIOR APPLICATION NUMBER: US 60/398,761
PRIOR FILING DATE: 2002-07-29
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 2301

APPLICANT: CHANG, Chiwen
TITLE OF INVENTION: USE OF SOLUBLE CD26 AS INHIBITOR OF
FILE REFERENCE: 39533-0001
CURRENT APPLICATION NUMBER: US/11/208,288
CURRENT FILING DATE: 2005-08-18
PRIOR APPLICATION NUMBER: US 60/605,013
PRIOR FILING DATE: 2004-08-26
NUMBER OF SEQ ID NOS: 8
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 2217
TYPE: DNA
ORGANISM: Homo sapiens
US-11-208-288-3

Query Match 2.5%; Score 77.8; DB 11; Length 2217;
Best Local Similarity 48.7%; Pred. No. 2.6e-10;
Matches 242; Conservative 0; Mismatches 252; Indels 3; Gaps 1;

QY 2270 TAGACAAACAGGGGATCCTGTCAACGAGGCTTAAATTTGAAGGCGCTTTAAATATATAAA 2329
DB 1649 TTGATGCGCAGAGGAGTGGTTACCAAGGAGATAAGATCATGTCATCAATCAACAGAGAC 1708
QY 2330 TGGGTCAATAGAAATTTGACGATCAGGTGGAGGACTCCATATCTAGCTTCTCGATATG 2389
DB 1709 TGGGAACATTTGAAGTTGAAGATCAAAATTTGAAGGAGCCAGACAATTTTCAAAAATGG 1765
QY 2390 ATTTCATTGACTTGTAGATCTGTGGGATCCAGCGCTGTCTATGGAGGATACCTCTCCC 2449
DB 1766 GATTTGTGGAACAAACAAATTTGCAATTTGGGGCTGTCTATGGAGGATACGTAACCT 1825
QY 2450 TGATGCGATTAAATGACAGAGTTCAGATATCTTCAGGGTTGCTTATGCTGGGGCCCCAGTCA 2509
DB 1826 CAATGGTCTGGGATCGGGAAGTGGCGTGTCTCAAGTGTGGAATAGCGGTGGCGCTGTAT 1885
QY 2510 CTCTGTGGATCTTCTATATAGATACAGGATACAGGAAGCTTATATGGGTCAACCTGACAC 2569
DB 1886 CCGGTGGGAGTACTATGACTCAGTGTACACAGAACGTTTACATGGGTCTCCCAACTCCAG 1945
QY 2570 ATGACAGGGCTATTACTTAGGATCTGTGGCCATCGAGCAGAGAAAGTTCCCTCTGAAAC 2629
DB 1946 AAGACACCTTGACCATTAACAGAAATTCACAGTATGACAGAGCTGAAATTTTAAAC 2005
QY 2630 CAAATCGTTTACTGCTTACATGTTTCTCGATGAGAAATGTCCTATTTTGCACATACCA 2689
DB 2006 AAGTTGAGTACCTCTTATTCATGGAACAGCAGATGATACGTTTCACTTTTCAGCAGTCAG 2065
QY 2690 GTATATPACGTATTTTATGAGGGCTGGAAAGCCATATGATTTACAGATCTATCTCTC 2749
DB 2066 CTCAGATCTCAAAGCCCTGGTGCATGTTGGAGTGGATTTCCAGGCAATGTGGTATATCTG 2125
QY 2750 AGGAGACACACAGGATA 2766
DB 2126 ATGAAGACCATGGAATA 2142

RESULT 10
US-10-522-789-1
Sequence 1, Application US/10522789
Publication No. US20050260732A1
GENERAL INFORMATION:
APPLICANT: TANIABE SEIYAKU CO., LTD.
TITLE OF INVENTION: three-dimensional structure of dipeptidyl peptidase IV
FILE REFERENCE: 03-039-PCT
CURRENT APPLICATION NUMBER: US/10/522,789
CURRENT FILING DATE: 2005-01-28
PRIOR APPLICATION NUMBER: US 60/398,761
PRIOR FILING DATE: 2002-07-29
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 2301

APPLICANT: CHANG, Chiwen
; TITLE OF INVENTION: USE OF SOLUBLE CD26 AS INHIBITOR OF
; ; TITLE OF INVENTION: ANGIOGENESIS AND INFLAMMATION
; FILE REFERENCE: 39533-0001
; CURRENT APPLICATION NUMBER: US/11/208,288
; CURRENT FILING DATE: 2005-08-18
; PRIOR APPLICATION NUMBER: US 60/605,013
; PRIOR FILING DATE: 2004-08-26
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 2283
; TYPE: DNA
; ORGANISM: Mus musculus
US-11-208-288-5

Query Match 2.3%; Score 71; DB 11; Length 2283;
Best Local Similarity 49.9%; Pred. No. 2.1e-08;
Matches 207; Conservative 0; Mismatches 205; Indels 3; Gaps 1;
QY 2270 TAGACAAACAGGGGATCTGTCCACCGAGGGCTTAAATTTGAAGGCGCTTTAAATATAAAAA 2329
DB 1715 TTGACGGCAGAGGAGTGGTTACCAAGGAGATAGATCATGCATGCAATCAACAGAGAT 1774
QY 2330 TGGGTCAATAGAAATGACGATCAGGTGGAAGGACTCCAAATATCTAGCTTCTCGATATG 2389
DB 1775 TGGGAACACTGGAAGTGAAGATCAAAATGAAGCAGCC---AGGCAATTTGTAAAAATGG 1831
QY 2390 ATTTCATTGACTTACATGTTGGGCATCCAGGCTGGTCTTATGAGGATACCTCTCCC 2449
DB 1832 GATTTGTGATAGCAAGCGAGTTGCAATTTGGGGCTGGTCATATGGAGGTATGTAACT 1891
QY 2450 TGATGCGATTAATGACAGAGTCCAGATATCTTCAGGGTTGCTATTGCTGGGGCCCCAGTCA 2509
DB 1892 CAATGCTCTGGATCGGAAGTGGCGTTCAGTGGGAGATAGCTGTGGACCTGTGT 1951
QY 2510 CTCTGTGGATCTTCTATGATACAGGATACACGGAACGTTATATGGGTCACTCTGACACAGA 2569
DB 1952 CACGTGGGAGTACTATGACTCAGTGTACACAGAGCGTTACATGGGTCTCCCAATCCAG 2011
QY 2570 ATGAACAGGGCTATTACTTAGTATCTGTGGCATCGAAGCAGAAAGTTCCCTCTGAAC 2629
DB 2012 AAGACAACTTGACCAATTACAGGAACTCAACAGTCATGACGAGCAGTGAACATTTTAAAC 2071
QY 2630 CAATCGTTTACTGCTTCTACATGTTTCTGGATGAGATGTCATTTTGCACA 2684
DB 2072 AAGTTGAGTACCTCTTATTTCATGAACGGCAGATGATATGTTTCACITTCAGCA 2126

RESULT 14
US-11-079-463-2536
; Sequence 2536, Application US/11079463
; Publication No. US20060073161A1
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTERIOIDES FRA
; ; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: PATH00-03DIV2
; CURRENT APPLICATION NUMBER: US/11/079,463
; CURRENT FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/128,705
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: US 09/540,209
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 2536
; LENGTH: 1884
; TYPE: DNA
; ORGANISM: B. fragilis
US-11-079-463-2536
Query Match 2.2%; Score 68.6; DB 11; Length 1884;
Best Local Similarity 48.2%; Pred. No. 8.5e-08;
Matches 184; Conservative 0; Mismatches 182; Indels 3; Gaps 1;

Matches 225; Conservative 0; Mismatches 239; Indels 3; Gaps 1;
QY 2119 AAGCCTCATGATCTACAGCTCGAAAGAAATATCTACTGTGTTTCATATATGGTGGT 2178
DB 1186 AAACCGGCGGATTTGATCCGAAACAAAGAAATACCGGCTATCGTATACGTATACGGCGT 1245
QY 2179 CTTACAGTGCAGTTGGTGAATTAATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACC 2238
DB 1246 CTTACAGCACAACCTGGTCAACACGGCTGGCAGAACGGTGCACGAGGTGGGATATCTAT 1305
QY 2239 CTAGCCTCTCTAGTTTATGTGTGTAGTGTAGATAGACAAACAGGGGATCTGTCAACCGAGGG 2298
DB 1306 ATGCCCAACAAAGGTTACATCATGTTTACCGTAGAGCGGACGTGGAGCAGCAATCGGGA 1365
QY 2299 CTTAAATTTGAAGGCGCTTTAAATATAAAATGAGTGGTCAAAATAGAAAATTTGACGATCAGGTG 2358
DB 1366 CTCGATTTTGAGAAATGTTACTTTTCCGCCCAATTCGAAAGCTCCCTACGTCGATGGGA---ACCGTATCGGAGTA 1425
QY 2359 GAAGGACTCCAATATCTAGCTTCTCGATATGATTTTCAATGACTTTAGATCGTGTGGCATC 2418
DB 1426 AAGGAAACCGAAATTCCTGAAAGCTCCCTACGTCGATGGGA---ACCGTATCGGAGTA 1482
QY 2419 CACGGCTGGTCTTATGGAGGATACCTCTCCCTGATGGCAATTAATCGCAGAGTCAAGATATC 2478
DB 1483 CACGGCTGGAGCTTTGGCGGTCAATGACCACTGCGCTTCTCCTTCGCTATCCGAGATA 1542
QY 2479 TTACAGGTTGCTTATGCTCGGGCCCCAGTCACTCTGTGATGCTTCTATGATACAGGATAC 2538
DB 1543 TTTAAAGTGGCGTGGCGGGCTCCCGTCAATTGACTGGGGTATTACGAAGTGTATGTAC 1602
QY 2539 ACAGGAGTTATATGGGTCACTCTGACCAAGATGAACAGGGCTATTA 2585
DB 1603 GGAGAACGCTATATGATATACCCCGCAAGCAATCCGAAAGGATATAA 1649

RESULT 15
US-11-245-147-168
; Sequence 168, Application US/11245147
; Publication No. US20060030541A1
; GENERAL INFORMATION:
; APPLICANT: GARCIA, TERESA
; APPLICANT: ROMAN ROMAN, SERGIO
; APPLICANT: BARON, ROLAND
; APPLICANT: CALL, KATHERINE
; APPLICANT: THEILHABER, JOACHIM
; APPLICANT: CONNOLLY, TIMOTHY
; APPLICANT: JACKSON, AMANDA
; APPLICANT: BUSHNELL, STEVEN
; APPLICANT: RAWADI, GEORGES
; TITLE OF INVENTION: GENES INVOLVED IN OSTEOGENESIS, AND METHODS OF USE
; FILE REFERENCE: 37991-0023
; CURRENT APPLICATION NUMBER: US/11/245,147
; CURRENT FILING DATE: 2005-10-07
; PRIOR APPLICATION NUMBER: PCT/IB02/02211
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: 60/281,400
; PRIOR FILING DATE: 2001-04-05
; NUMBER OF SEQ ID NOS: 246
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 168
; LENGTH: 2814
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Homo sapiens fibroblast activation protein, alpha
; OTHER INFORMATION: (FAP), mRNA
US-11-245-147-168
Query Match 2.1%; Score 65.8; DB 11; Length 2814;
Best Local Similarity 46.8%; Pred. No. 6.5e-07;
Matches 242; Conservative 0; Mismatches 272; Indels 3; Gaps 1;

QY 2250 AGGTTATGTGTTGTAGTATAGAACAGGGGATCCTGTCCAGCGGCTTAAATTTGA 2309

Db	1903	AGGATGTCATTCGCTTGGTGGATGTCGAGGACAGCTTCCAGGTGACAACTCCT	1962
Qy	2310	AGGCCCTTTAAATATATAATAGGTCAAATAGAAATTGACGATCAGGTGGAGGACTCCA	2369
Db	1963	CTATCCAGTGTATCGAAAGCTGGGTGTTTATGAAGTTGAAGACCAGATTACAGCTGTGAG	2022
Qy	2370	ATATCTAGCTTCTCGATATGATTTTCATGTAGTCTAGATCGTGTGGGCAITCCA	2429
Db	2023	AAA---ATTCTAGAAATGGGTTCATTTGATGAAAAAAGAAATAGCCATATGGGGCTGGTC	2079
Qy	2430	CTATGGAGGATACCTCTCCCTGATGGCAITTAATGCGAGGTCAGATATCTTCAGGGTTGC	2489
Db	2080	CTATGGAGGATACGTTTCATCACTGGCCCTTGCAITCTGGAATCTGTTTCAAATGTGG	2139
Qy	2490	TATTGCTGGGGCCCGAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACGTTA	2549
Db	2140	TATAGCAGTGGTCCAGTCTCCAGTGGGAATATTACGCGTCTGTCTACACAGAGAGATT	2199
Qy	2550	TATGGTCACCTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGCCCATGCAAGC	2609
Db	2200	CATGGTCTCCCAACAAAGGATGATAATCTTGACACTATAAGAAATCAACTGTGATGCG	2259
Qy	2610	AGAAAGTTCCCTCTGAACCAATCGTTTACTGCTCTTACATGTTTCTCTGGATGAGAA	2669
Db	2260	AAGCAGAAATTTTCAGAAATGTAGACTATCTCTCATCCAGGAACAGCAGATGATAA	2319
Qy	2670	TGTCCATTTTGACATACCACTATATTACTGAGTTTTTTAGTGAGGGCTGGAAAGCCATA	2729
Db	2320	TGTGCACCTTCAAAACCTCAGCAGATGCTAAGCTCTGGTTATGTCACAAAGTGGATTT	2379
Qy	2730	TGATTTACAGATCTATCTCTCAGGAGAGACACAGCATTA	2766
Db	2380	CCAGGCAATGTGGTACTCTGACCAAGAACCAAGGCTTA	2416

Search completed: April 15, 2006, 01:59:35
Job time : 1181 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: April 14, 2006, 10:51:01 ; Search time 90.7 Seconds
(without alignments)
3022.849 Million cell updates/sec

Title: US-10-825-632-2

Perfect score: 5552

Sequence: 1 aagtgttaagctcgagg.....agaattactaaaaaaaaa 3120

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 4886326

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+n2p.model -DEV=xlbp
-O=/abss/ABSSWEB_epool/US10825632/runat_14042006_092348_3416/app_query.fasta.1
-DB=A Geneseq -QFMT=fastan -SUFFIX=n2p.rag -MINMATCH=0.1 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss02p
-USER=US10825632.ecgn_1_1_348 @runat_14042006_092348_3416 -NCPU=6 -ICPU=3
-NO MWAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

A_Geneseq_21:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*
9: Geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4700	84.7	882	4 AAB47187	Aab47187 Human DPP
2	4700	84.7	882	5 ABG61591	Abg61591 Human DPP
3	4700	84.7	882	5 AAG78415	Aag78415 Amino aci
4	4700	84.7	882	5 AAE24170	Aae24170 Human dip
5	4700	84.7	882	5 AAU74749	Aau74749 Human pro
6	4700	84.7	882	5 ADI17086	Adi17086 Human NOV
7	4700	84.7	882	6 ABU07720	Abu07720 Human ser
8	4528.5	81.6	883	5 ADI17085	Adi17085 Murine NO
9	4397.5	79.2	831	6 ABU92026	Abu92026 Human pro

10	4353	78.4	824	6 ABU92030	Abu92030 Human pro
11	4118	74.2	782	5 ABB97361	Abb97361 Novel hum
12	3901	70.3	746	8 ADI16334	Adi16334 Human pro
13	3817	68.8	738	8 ADT04044	Adt04044 Human pro
14	3771	67.9	724	5 ABB97362	Abb97362 Novel hum
15	3607.5	65.0	690	5 ABG61594	Abg61594 Human DPR
16	3513.5	63.3	661	5 ABG61596	Abg61596 Human DPR
17	3504	63.1	658	5 ABG61600	Abg61600 Human DPR
18	3333.5	60.0	632	4 AAB93565	Aab93565 Human pro
19	3236	58.3	613	5 ABG61601	Abg61601 Human DPR
20	3087	55.6	587	5 ADI13398	Adi13398 Human CD-
21	2871	51.7	892	5 ABG61602	Abg61602 Human DPR
22	2871	51.7	892	5 ABG61604	Abg61604 Human DPR
23	2871	51.7	892	5 ABB98134	Abb98134 Human PMM
24	2871	51.7	892	8 ADS10951	Ads10951 Human the
25	2871	51.7	863	5 AAE24168	Aae24168 Human dip
26	2870	51.7	863	5 ABG61592	Abg61592 Human DPP
27	2870	51.7	863	5 ADI17083	Adi17083 Human NOV
28	2862	51.5	863	5 ADI16690	Adi16690 Human NOV
29	2862	51.5	863	5 ADI16688	Adi16688 Human NOV
30	2862	51.5	863	8 ADN42344	Adn42344 Human nov
31	2852.5	51.4	580	5 AAE14337	Aae14337 Human pro
32	2842	51.2	863	8 ADN42342	Adn42342 Human nov
33	2835	51.1	830	5 AAE24171	Aae24171 Human dip
34	2833	51.0	869	5 AAE24169	Aae24169 Alternati
35	2821.5	50.8	879	5 ABG61607	Abg61607 Human DPR
36	2821.5	50.8	879	5 ABG61608	Abg61608 Human DPR
37	2821.5	50.8	879	8 ADS10953	Ads10953 Human the
38	2713.5	48.9	847	5 AAE23875	Aae23875 Murine di
39	2708	48.8	854	8 ADS10952	Ads10952 Human the
40	2672.5	48.1	873	8 ADS10479	Ads10479 Human the
41	2523	45.4	493	7 ADE78977	Ade78977 Human pro
42	2505	45.1	755	8 ADQ67811	Adq67811 Novel hum
43	2470	44.5	516	6 ABU92029	Abu92029 Human pro
44	2422	43.6	465	4 AAB47189	Aab47189 Human DPP
45	2413.5	43.5	737	4 AAM38724	Aam38724 Human pol

ALIGNMENTS

RESULT 1

AAB47187
ID AAB47187 standard; protein; 882 AA.

XX AAB47187;

DT 29-JUN-2001 (first entry)

XX Human DPP8.

XX Human; dipeptidyl aminopeptidase; DPP8; prollyl oligopeptidase;
dipeptidyl peptidase; DPPIV; T cell; cleavage; diarrhoea;
growth hormone deficiency; glucose level; mucosal regeneration;
non-insulin dependent diabetes mellitus; glucose intolerance;
immunosuppression.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Active-site 739 /note= "Forms part of Ser-Asp-His catalytic triad"

FT Active-site 817 /note= "Forms part of Ser-Asp-His catalytic triad"

FT Active-site 849 /note= "Forms part of Ser-Asp-His catalytic triad"

FT WO200119866-A1.

XX 22-MAR-2001.

XX 11-SEP-2000; 2000WO-AU001085.

XX 10-SEP-1999; 99AU-00002762.

PR 18-FEB-2000; 2000AU-00005709.
 XX (UNSY) UNIV SYDNEY.
 XX PA Abbott CA, Gorell MD;
 XX PI WPI; 2001-281520/29.
 XX DR N-FSDB; AAC8594.
 XX New human dipeptidyl aminopeptidase (Dpp8) useful for cleaving
 PT substrates, identifying inhibitors of dpp8 catalytic activity which have
 PT therapeutic uses, and for detecting activated T cells.
 XX Claim 1; Fig 2; 78pp; English.
 XX This sequence represents human dipeptidyl aminopeptidase (Dpp8). Dpp8 has
 CC substrate specificity for H-Ala-Pro-pNA, H-Gly-Pro-pNA and H-Arg-Pro-pNA.
 CC therefore, it is a prolyl oligopeptidase and a dipeptidyl peptidase,
 CC because it is capable of hydrolysing the peptide bond C-terminal to Pro
 CC in each of these compounds. Dpp8 is homologous with human DPP4. Dpp8 is
 CC useful for cleaving a substrate, and for detecting an activated T cell.
 CC which involves measuring the level of DPP8 gene expression in a T cell.
 CC The level of DPP8 expression is detected by identifying a molecule capable of
 CC inhibiting the cleavage of the substrate by DPP8. Molecules identified as
 CC inhibiting DPP8 catalytic activity may be useful for treating diarrhoea,
 CC growth hormone deficiency, lowering glucose levels in non-insulin
 CC dependent diabetes mellitus and other disorders involving glucose
 CC intolerance, enhancing mucosal regeneration and as immunosuppressants
 XX Sequence 882 AA;

Alignment Scores:
 Pred. No.: 0 Length: 882
 Score: 4700.00 Matches: 882
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 84.7% Indels: 0
 DB: Gaps: 0

US-10-825-632-2 (1-3120) x AAB47187 (1-882)

QY 214 ATGGCAGCAGCAATGAAACAGACAGCTGGTGTGAGATATTTGAACTCGGACTGT 273
 Db 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
 QY 274 GAGGAGATATTAATACAGGATCGGCTAAATGGAGCCTTTTATGTGACCGTAT 333
 Db 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
 QY 334 TCCTGGAGTCAGCTTAAAGCTGCTTCCGATACAGAAATATCATGCGTCATGATG 393
 Db 41 SerTrpSerGlnLeuLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
 QY 394 GCTAAGGCACCATGATTTTCATGTTGTGAAGAGAAATGATCCAGATGGACCTCATTTCA 453
 Db 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
 QY 454 GACAGATCTATTACCTTCGATGCTGGTGAAGAGAAATACACTGTTTATTTCT 513
 Db 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
 QY 514 GAATTCGCCAACTATCAATAGCAGCAGCTTAAATGCTCTCTTGGAGCCCTTTTGG 573
 Db 101 GluIleProLysThrIleAsnArgAlaValLeuMetLeuSerTrpLysProLeuLeu 120
 QY 574 GATCTTTTCGGCACACTGGATGATGAAATGATTTCTCCGAGAGAGAACTATTAAAGA 633
 Db 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140
 QY 634 GAAAGAAAACGATTTGGACAGTCGGAATGCTTCTTACGATATACCAAGGAAGTGA 693
 Db 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160

QY 694 ACATTTCTGTTTCAAGCCGGTAGTGGAAATTTATCAGTAAAGATCGAGCCGCAAGGA 753
 Db 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyProGlnGly 180
 QY 754 TTTACGCAACACCTTTAAAGCCCAATCTAGTGGAAACTAGTTGTCCCAACATACCGATG 813
 Db 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
 QY 814 GATCCAAAATATGATCCCGCTGATCCAGACTGGATGCTTTTATATACATAGCAACGATATT 873
 Db 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
 QY 874 TGGATATCTAACATCGTAACCCAGAGAGAAAGAGACTCATTATGTGCAATAGCTA 933
 Db 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu 240
 QY 934 GCCACATGGAAGAAGATGCCAGATCAGCTGGAGTCCCTACCTTGTCTCCCAAGAGAA 993
 Db 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu 260
 QY 994 TTTGATAGATATCTGGCTATTGGTGTGTCCTCAAAAGCTGAAACAACTCCACGTGGTGGT 1053
 Db 261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280
 QY 1054 AAAATCTTAGAATTTCTATATGAAAGAAATGATCAATCTGAGGTGGAAATTTATTCATGTT 1113
 Db 281 LysIleLeuArgIleLeuTyrGluGluAsnAspLysSerGluValGluIleHisVal 300
 QY 1114 ACATCCCTATGTTGGAAACAGAGGAGGAGCAGATTCCGTTATCTTAAACAGGATACA 1173
 Db 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320
 QY 1174 GCAATCTCTAAAGTCACCTTTTAAGATGTCAGAAATATATGATGATGCTGAAGAGAGATC 1233
 Db 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
 QY 1234 ATAGATGTCATAGATAAGGAACATAATTCACCTTTTGGATCTTATTTGAAGAGATTGAA 1293
 Db 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
 QY 1294 TATATTGCCAGAGCTGGATGGACTCTCGAGGAGAAATATGCTCGTCCATCTACTAGAT 1353
 Db 361 TyrIleAlaAlaGlyTyrThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
 QY 1354 CGCTCCACAGACTCGCTACAGATAGTGTGTGATCTCACCTGAATTTATTTATCCCACTAGAA 1413
 Db 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
 QY 1414 GATGATGTTATGGAAGGAGAGACTCATTGAGTCAGTCCCTGATTTCTGTGACGCCACTA 1473
 Db 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
 QY 1474 ATTATCTATGAGAAACACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTTTT 1533
 Db 421 IleIleTyrGluGluThrThrAspIleThrIleAsnIleHisAspIlePheHisValPhe 440
 QY 1534 CCCCAGAGTCACGAGAGAAATGAGTTTATTTTGGCTCTGATGCAAAACAGGTTTC 1593
 Db 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
 QY 1594 CGCTATTATACAAAATTACATCTATTTTAAAGGAAAGCAATATATAACGATCCAGTGT 1653
 Db 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
 QY 1654 GGGCTGCTCTCAAGTGAATTTCAAGTCTCTATCAAGAGAGAGATAGCAATTTACCATG 1713
 Db 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 500
 QY 1714 GGTCAATGGGAAGTTCTTGGCCGGCATGGATCTTAATATCCAAAGTTGATGAAGTCAGAGG 1773
 Db 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520

```
QY 1774 CTGGTATATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCACGTGACGTAGTCAGT 1833
Db 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer 540
QY 1834 TAGTAAATCCCTGGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACATTCTTGCTGC 1893
Db 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
QY 1894 ATCAGTCAGCACTGTGACTTCTTTATAAGTAAGTATAGTAACCAAGAAGATCCACACTGT 1953
Db 561 IleSerGlnHisCysAspPheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
QY 1954 GTGTCCTTTACAGCTATCAAGTCCTCAAGATGACCCCACTTGCAAAAACAAGAAATTT 2013
Db 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
QY 2014 TGGGGCACCAATTTGGATTACGACGCTCTCTCTGACTATATCTCTCAGAGAAATTTTC 2073
Db 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePhe 620
QY 2074 TCTTTTGAAGTACTGCTGATTTACATTGTATGGATGCTCTACAAGCCTCATGATCTA 2133
Db 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
QY 2134 CAGCTGGAAGAAATATCCTACTCTGCTGTTCATATATATGCTGCTCAGGTGCGAGTTG 2193
Db 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 660
QY 2194 GTGAATATCGTTTAAAGGAGTCAAGTATTTCCGCTTGATACCTAGCTCTCTAGGT 2253
Db 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680
QY 2254 TATGTGCTGTAGTATACAAACAGGGGATCCTGCTACCGAGGGCTTAAATTTGAAGGC 2313
Db 681 TyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 700
QY 2314 GCCTTTAAATATAAAATGGGTCAAAATAGAAATTCAGCATCAGGTGGAAGACTCCCAAT 2373
Db 701 AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr 720
QY 2374 CTAGCTTCGATATGATTTCACTTCACTTAGATCGTGGGCATCCACGCTGCTCTAT 2433
Db 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyr 740
QY 2434 GGAGGATACCTCTCCCTGATGCATTAAATGCAGAGCTCAGATATCTTCAGGTGCTATT 2493
Db 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
QY 2494 GCTGGGGCCCCAGTCACTCTGGGATCTTCTATGATACAGATACACGGAACGTTATATG 2553
Db 761 AlaGlyAlaProValThrLeuThrIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780
QY 2554 GGTCACCTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCGAA 2613
Db 781 GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
QY 2614 AAGTCCCTCTGAACCAATCGTTTACTGCTCTACATGCTTTCCTGGATGAGAATGTC 2673
Db 801 LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
QY 2674 CATTTTGCACATACCACTATATTACTAGTTTTTTAGTGAGGGCTGGAAAGCCCATATGAT 2733
Db 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
QY 2734 TTACAGATCTATCTCAGGAGACACAGCAATAGAGTTCTCTGAATCGGAGAACATTAT 2793
Db 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
QY 2794 GAATGCACTCTTTGGCACTACCTTCAGAAAACCTTGGATCAGCTATTGCTGCTCAAAA 2853
Db 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLys 880
QY 2854 GTGATA 2859
```

```
Db 881 ValIle 882
RESULT 2
ABG61591
ID ABG61591 standard; protein; 882 AA.
AC ABG61591;
XX 12-AUG-2002 (first entry)
XX Human DPPIV related serine protease DPRP-1.
XX Human; serine protease; dipeptidyl peptidase IV-related protein; DPRP;
KW DPPIV; infection; human immunodeficiency virus; HIV-1; HIV-2; pain;
KW diabetes; infertility; obesity; anorexia; Parkinson's disease; stroke;
KW heart failure; hypertension; urinary retention; osteoporosis; cancer;
KW ulcer; allergy; cancer; psychotic disorder; neurological disorder;
KW dyskinesia; reproductive disorder; inflammatory disorder;
KW metabolic disorder.
XX Homo sapiens.
XX WO200231134-A2.
XX 18-APR-2002.
XX 12-OCT-2001; 2001WO-US031874.
XX 12-OCT-2000; 2000US-0240117P.
XX (FERR ) FERRING BV.
XX Qi S, Akineanya KO, Riviere EJ, Junien J;
XX WPI, 2002-444178/47.
XX N-PSDB; ABK83322.
XX New dipeptidyl peptidase IV-related proteins and nucleic acids encoding
PT the proteins, useful for treating e.g. fungal, bacterial, protozoan and
PT viral infections, cancers, allergies, neurological disorders, or pain.
XX Claim 17; Fig 1; 113pp; English.
XX The present invention relates to the isolation of novel human serine
CC proteases referred to as dipeptidyl peptidase IV (DPPIV)-related proteins
CC (DPRP). The dipeptidyl peptidase IV-related proteins (DPRP) and nucleic
CC acids encoding them are useful for treating infections such as fungal,
CC bacterial, protozoan and viral infections, particularly infections caused
CC by human immunodeficiency virus (HIV-1 or HIV-2), pain, diabetes,
CC precocious puberty, infertility, obesity, anorexia, bulimia, Parkinson's
CC disease, acute heart failure, hypotension, hypertension, urinary
CC retention, osteoporosis, angina pectoris, stroke, ulcers, asthma,
CC allergies, cancers, migraine, vomiting, psychotic and neurological
CC disorders (e.g. anxiety, dementia, or schizophrenia), and dyskinesias.
CC These may also be used in discovering therapeutic agents for the
CC treatment of reproductive, inflammatory and metabolic disorders.
XX -ABG61591 represent human DPRP proteins
XX Sequence 882 AA;
SQ
Alignment Scores:
Pred. No.: 0 Length: 882
Score: 4700.00 Matches: 882
Percent Similarity: 100.0% Conservatives: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 84.7% Indels: 0
DB: 5 Gaps: 0
US-10-825-632-2 (1-3120) x ABG61591 (1-882)
QY 214 ATGGCAGCAGCAATGGAAACAGAACAGCTGGGTGTGAGATATTTGAACCTCGGACTGT 273
|||||
```

Db	1	MetAlaAlaMetGluThrGluGlnLeuGlyValGluLeuPheGluThrAlaAspCys	20
Qy	274	GAGGAGAAATTTGAATCAAGATCGCGCTAAATTTGAGGCTTTTATGTTAGCGGTAT	333
Db	21	GluGluAsnIleGlnSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr	40
Qy	334	TCCTGGAGTCAGCTTTAAAGCTGCTTGGCCGATACCCAGAAATATATCATGCTACATGATG	393
Db	41	SerTrpSerGlnLeuLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet	60
Qy	394	GCTAAGGCCACCATGATTTTCATGTTGTGAAGAGGAATGATCCAGATGCACCTCATCA	453
Db	61	AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer	80
Qy	454	GACAGAAATCTATTACCTTGCCTATGCTGCTGAGAACAGAGAAATACACTGTTTATTCT	513
Db	81	AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer	100
Qy	514	GAATTTCCCAAACTATCAATAGAGCAGCAGTCTTAATGCTCTCTTGGAGCCCTCTTTTG	573
Db	101	GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu	120
Qy	574	GATCTTTTTCAGGCAACACTGACATGGAATGTATTTCTCGAAGAAGAACTATTAAAG	633
Db	121	AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg	140
Qy	634	GAAGAAGAACGATGGAACAGTCGGAATTCCTCTTACGATTTATCACCAAGAGAGTGA	693
Db	141	GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly	160
Qy	694	ACATTTCTTTTCAAGCCGCTAGTGGAAATTTATCACGTAAAGATGAGGGCCACAAAGA	753
Db	161	ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly	180
Qy	754	TTTACGCACAACCTTTTAAGGCCAATCTAGTGGAAACTAGTTGTGCCAACAATCCGATG	813
Db	181	PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet	200
Qy	814	GATCCAAAATTTATGCCCGCTGATCCAGACTGGATGCTTTTATACATAGCAACGATATT	873
Db	201	AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle	220
Qy	874	TGATATCTAAACATCGTAACACAGAGAAGAAAGAGACTCACTATGTGCAACAATGAGCTA	933
Db	221	TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu	240
Qy	934	GCCAAATGGAAGAGATCCAGATCAGCTGGAGTCGCTACTTTGTTCTCCAAAGAGAA	993
Db	241	AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu	260
Qy	994	TTTGTAGATATTCTGGCTATTGTTGGTGTCCAAAGCTGCAACAACTCCCACTGGTGTG	1053
Db	261	PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrTrpSerGlyGly	280
Qy	1054	AAAATTTAGAAATCTATACAGAAATGATGAACTGAGTGGGAATTTATTCATGTT	1113
Db	281	LysIleLeuArgIleLeuTyrGluAsnAspGluSerGluValGluIleIleHisVal	300
Qy	1114	ACATCCCTTATGTTGAAACAAAGGAGGCGAGATTCATTCCTGTTATCTTAAACAGGTACA	1173
Db	301	ThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThr	320
Qy	1174	GCAATCTCTAAAGTCACTTTTAAAGATGTCAGAAATATATGATGCTGAAGGAGATC	1233
Db	321	AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle	340
Qy	1234	ATAGATGTCATAGATAAGGAACATAATTCACCTTTTGAGATTCATTGGAAGAGTTGAA	1293
Db	341	IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu	360
Qy	1294	TATATTGCCAGACTGGATGGACTCTGAGGGAAATATGCTTGGTCCATCCTACTAGAT	1353
Db	361	TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaIleTrpSerIleLeuLeuAsp	380
Qy	1354	CGCTCCAGACTCGCTACAGATAGTTGATCTCACCTGAATTTATTTATCCCGTAGAA	1413
Db	381	ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu	400
Qy	1414	GATGATGTTATGGAAGCAGAGACTCAATGAGTCAGTCAGTCGCTGATTTCTGTGAGCCACTA	1473
Db	401	AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu	420
Qy	1474	ATTATCTATGAAGAAAACAACAGACATCTGGATAAATATCCATGCATCTTCTCATGTTT	1533
Db	421	IleIleTyrGluGluThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe	440
Qy	1534	CCCCAAGTCACGAAGAGAAATTGAGTTTATTTTGGCTCTGAATCCAAAACAGGTTTC	1593
Db	441	ProGlnSerHisGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe	460
Qy	1594	CGTCATTTATCAAAAATACATCTATTTTAAAGGAAACAAATAATAACATCCAGTGGT	1653
Db	461	ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly	480
Qy	1654	GGCTCGCTCCCAAGTGATTTCAAGTGTCTCTCAAGAGGAGAGATAGCAATTACCAGT	1713
Db	481	GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer	500
Qy	1714	GGTGAATGGAAAGTTCTTGGCCGATCGATCTAATATCCAAAGTTGATGAAGTCAGAAAG	1773
Db	501	GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg	520
Qy	1774	CTGGTATATTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTAGTGTAGT	1833
Db	521	LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisIleLeuTyrValValSer	540
Qy	1834	TAGTAAATCTCGGAGAGTGCACAGGCTGACTGACCGTGCCTACTCACATCTTCTGCTGC	1893
Db	541	TyrValAsnProGlyLysValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys	560
Qy	1894	ATCAGTCAGACTGTGACTTCTTTTAAAGTAAAGTATAGTAACAGAGAATCCCACTGT	1953
Db	561	IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys	580
Qy	1954	GTCTCCCTTTACAGCTATCAAGTCTCAAGATGACCCCACTTCGCAAAAACAAAGAAATTT	2013
Db	581	ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe	600
Qy	2014	TGGGCCACCAATTTTGGATTCAGCAGGTCTCTTCTGCTGACTATATCTCTCCAGAAATTTTC	2073
Db	601	TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePhe	620
Qy	2074	TCTTTGAAAGTACTACTGATTTACATTTGATGGATGCTCTCAAGCCCTCATGATCTA	2133
Db	621	SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu	640
Qy	2134	CAGCTGGAAGAAATATCTACTGTGCTGTTTCATATATGTTGCTGCTCAGGTGCAGTTG	2193
Db	641	GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu	660
Qy	2194	GTGAATAATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCTAGCCTCTCTAGGT	2253
Db	661	ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly	680
Qy	2254	TATGTGCTGTAGTATAGACACACAGGGGATCCTCTGACCCGAGGCTTTAAATTTGAAAGC	2313
Db	681	TyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly	700
Qy	2314	GCCTTTAAATATTAATGGTCAATAGAAATGCAATGAGTACAGGTGGAGGACTCCAATAT	2373
Db	701	AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr	720
Qy	2374	CTAGCTTCTCGATATGATTTTCATTGACTTAGATTCGTGTGGCATCCAGGCTGCTCAT	2433
Db	721	LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyr	740

```
QY 2434 GGAGGATACCTCTCCCTGATGGCATTAATGACAGAGGTACAGATATCTTCAGGGTGTCTATT 2493
DB 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
QY 2494 GCTGGGGCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGACGTATATG 2553
DB 761 AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780
QY 2554 GGTCACTCCCTCAGCAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAA 2613
DB 781 GlyHisProAspGlnHisGlnGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
QY 2614 AAGTTCCTCTGACCAATCGTTTACTGCTCTTACATGTTCTCTGATGAGAATGTC 2673
DB 801 LysPheProSerGluProAsnArgLeuLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
QY 2674 CATTTGCACATACAGTATATATCTAGTGTATTTTACTGAGGCTGGAAAGCCATATGAT 2733
DB 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
QY 2734 TTACAGATCTATCTCAGGAGACACACAGCATAAAGAGTTTCTGTAATCGGAGACATTAT 2793
DB 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
QY 2794 GAACTGCATCTTTTGCACACTACCTTCAAGAAACCTTGGATCACGTAATGCTGCTTAAAA 2853
DB 861 GluLeuHisLeuLeuHisTyrLeuGlnGlnLeuLeuGlySerArgIleAlaAlaLeuLys 880
QY 2854 GTGATA 2859
DB 881 ValIle 882

RESULT 3
ID AAG78415 standard; protein; 882 AA.
XX
AC AAG78415;
XX
DT 12-APR-2002 (first entry)
XX
DE Amino acid sequence of 21953 human prollyl oligopeptidase.
XX
KW 21953 prollyl oligopeptidase; antibody; proline; endopeptidase; cancer;
KW cardiovascular disease; autoimmune disease; atopic allergy;
KW neuronal disorder; vascular disorder; prostate disorder; Cytostatic;
KW antidiabetic; antiarthritic; antiasthmatic; antinflammatory;
KW diabetes mellitus; arthritis; multiple sclerosis; asthma;
KW Grave's disease; neuronal disorder; demyelinating disease.
XX
OS Homo sapiens.
XX
PN WO200179473-A2.
XX
PD 25-OCT-2001.
XX
PF 11-APR-2001; 2001WO-US040483.
XX
PR 18-APR-2000; 2000US-0197508P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Meyers RA, Williamson M;
XX
DR WPI; 2002-034353/04.
XX
DR N-PSDB; AAH99934.
XX
PT New polypeptides 21953, member of human prollyl oligopeptidase family,
PT useful as diagnostic targets and therapeutic agents for controlling
PT cancer, lymphoma and leukemia.
XX
PS Claim 1; Page 102-103; 12ipp; English.
XX
CC This invention relates to an isolated 21953 human prollyl oligopeptidase.
```

```
CC Which is cytostatic, antidiabetic, antiarthritic, neuroprotective,
CC antithyroid, dermatological, antipsoriatic, antiasthmatic,
CC ophthalmological, antinflammatory, nootropic, antiparkinsonian,
CC anticonvulsant, gynaecological, vasotropic, antianginal, cardiant,
CC antiatherosclerotic, anorectic and metabolic in its action. Uses include
CC gene therapy, expression or activity of 21953 protein modulator, it is
CC useful for identifying a compound which binds to it and can be used in
CC preventing, treating or detecting a cellular proliferative or
CC differentiative disorder. The 21953 molecules can act as novel diagnostic
CC targets and therapeutic agents for controlling disorders associated with
CC the aberrant activity or degradation of peptide hormones e.g., disorders
CC associated with cell differentiation and proliferation such as cancer,
CC immune function, reproductive, neurological and cardiovascular function.
CC The 21953 molecules are thus useful for treating and preventing cellular
CC proliferative and differentiative disorders, haematopoietic neoplastic
CC disorders, immune disorders such as autoimmune diseases, diabetes
CC mellitus, arthritis, multiple sclerosis, asthma, Grave's disease,
CC neuronal disorders, demyelinating diseases, vascular disorders and
CC metabolism or pain disorders. This sequence represents the amino acid
CC sequence of 21953 human prollyl oligopeptidase
XX
SQ Sequence 882 AA;
```

```
Alignment Scores:
Pred. No.: 0 Length: 882
Score: 4700.00 Matches: 882
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 84.7% Indels: 0
DB: 5 Gaps: 0

US-10-825-632-2 (1-3120) x AAG78415 (1-882)

QY 214 ATGCAGCAGCAATGGAACAGAACAGCTGGGTGTGTAGATATTTGAACTCGGACTGT 273
DB 1 MetAlaAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
QY 274 GAGGAGAATATTGAATCACAGGATCGGCTTAATGGAGCCCTTTTATGTGGCGGTAT 333
DB 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
QY 334 TCCTGGAGTCACTTAAAGCTGCTGGCATACCAAGAAATATCATGGTACATGATG 393
DB 41 SerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
QY 394 GCTAAGGCACCATGATTTTCATGTTTGAAGAGGAATGATCCAGATGGACCTCATTTCA 453
DB 61 AlalysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
QY 454 GACAGAATCTATTACCTTGCCTATGTCGTGGAGAACAGAGAAATAACACTGTTTATTCT 513
DB 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
QY 514 GAAATTCCTCAAACTATCAATAGACGACGCTTAATGCTCTCTTGGAGCCTCTTTTG 573
DB 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
QY 574 GATCTTTTTCAGGCAACACTGGACTATGGAATGCTATTTCTCGAGAGAGAACTATTAA 633
DB 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetCysSerArgGluGluLeuLeuArg 140
QY 634 GAAAGAAAACGATTGGAAACAGTCGGAAATTCCTTACGATTATCACCAAGAGAGTGA 693
DB 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
QY 694 ACATTTCTGTTTCAAGCCGCTAGTGAATTTTATACGTAAGAGATGGAGGCCACACAGA 753
DB 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
QY 754 TTTAGGCAACACCTTTAAGCCCAATCTAGTGGAACTAGTTGTCCCAACATACGATG 813
DB 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
```

Which is cytostatic, antidiabetic, antiarthritic, neuroprotective, antithyroid, dermatological, antipsoriatic, antiasthmatic, ophthalmological, antinflammatory, nootropic, antiparkinsonian, anticonvulsant, gynaecological, vasotropic, antianginal, cardiant, antiatherosclerotic, anorectic and metabolic in its action. Uses include gene therapy, expression or activity of 21953 protein modulator, it is useful for identifying a compound which binds to it and can be used in preventing, treating or detecting a cellular proliferative or differentiative disorder. The 21953 molecules can act as novel diagnostic targets and therapeutic agents for controlling disorders associated with the aberrant activity or degradation of peptide hormones e.g., disorders associated with cell differentiation and proliferation such as cancer, immune function, reproductive, neurological and cardiovascular function. The 21953 molecules are thus useful for treating and preventing cellular proliferative and differentiative disorders, haematopoietic neoplastic disorders, immune disorders such as autoimmune diseases, diabetes mellitus, arthritis, multiple sclerosis, asthma, Grave's disease, neuronal disorders, demyelinating diseases, vascular disorders and metabolism or pain disorders. This sequence represents the amino acid sequence of 21953 human prollyl oligopeptidase

QY	814	GATCCAAATATATCCCGCTGATCCAGACTGGATGGCTTTTATATACATAGCAACGATATT	873	Db	561	IIeSerGlnHisCysAspPhePheIleSerLysTy-SerAsnGlnLysAsnProHisCys	580
Db	201	AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle	220	QY	1954	GTGTCCCTTTCAAGCTATCAAGTCTCTGAAGATGACCAACTTGGCAAAACAAGGAAATTT	2013
QY	874	TGGATATCTCAATCGTAACAGAGAGAGAGAGACTCACTTATGTGCACAACTGAGCTA	933	Db	581	ValSerLeuTyLysLeuSerProGluAspAspProThrCysLysThrLysGluPhe	600
Db	221	TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu	240	QY	2014	TGGGCCACCATTTTGGATTTCAGCAGGCTCTCTCTGACTATATCTCTCCAGAAATTTTC	2073
QY	934	GCCAACTGGAGAGATGCCAGATCAGCTGGATCGCTACCTTTCTTCCAAAGAGAA	993	Db	601	TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe	620
Db	241	AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu	260	QY	2074	TCTTTTGAAGTACTACTTGGATTTCATATTGTATGGGATGCTCTACAGGCTCATGATCTA	2133
QY	994	TTTGATAGATATCTCGGCTATTTGGTGTCTCCAAAGCTGAAACACTCCCACTGGTGT	1053	Db	621	SerPheGluSerThrThrGlyPheThrLeuTyLysMetLeuTyLysProHisAspLeu	640
Db	261	PheAspArgTySerGlyTyTrpTrpCysProLysAlaGluThrThrProSerGlyGly	280	QY	2134	CAGCTCGAAAGAAATATCTCTACTGTCTGTTCATATATATGCTGCTCTCAGGTGCAAGTTG	2193
QY	1054	AAAACTTTPAGAAATCTATATGAAGAAATGATGAAATCTGAGTGGAAATTTATCTGTT	1113	Db	641	GlnProGlyLysLysTyProThrValLeuPheIleTyGlyGlyProGlnValGlnLeu	660
Db	281	LysIleLeuArgIleLeuTyGluGluAsnAspGluSerGluValGluIleIleHisVal	300	QY	2194	GTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCCCTAGGCTCTCTAGGT	2253
QY	1114	ACATCCCTATGTTTGGAAACAAAGAGGCGCAGATTCATTCGTTATCTTAAACAGGTACA	1173	Db	661	ValAsnAsnArgPheLysGlyValLysTyPheArgLeuAsnThrLeuAlaSerLeuGly	680
Db	301	ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr	320	QY	2254	TATGTGTTGTAGTATGATAGACAAACAGGGGATCCTGTCAACGAGGCTTAAATTTGAAGGC	2313
QY	1174	GCAATCTCTAAAGTCACCTTTTAAAGATGTGAGAAATATGATTCATGCTGAAGAGATC	1233	Db	681	TyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly	700
Db	321	AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle	340	QY	2314	GCCTTTAAATATAAATCGGTCAATATAGAAATTCAGCATCAGGTGGAGGACTCCCAATAT	2373
QY	1234	ATAGATGTCATAGATAAGGAATTAATCAACCTTTTGAGATTCCTATTGGAAGAGTTGAA	1293	Db	701	AlaPheLysTyLysMetGlyGlnIleGluIleAspGlnValGluGlyLeuGlnTy	720
Db	341	IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu	360	QY	2374	CTAGCTTCGATATGATTTTCACTTAGATCGTGTGGCATCCAGGCTGTCTCTAT	2433
QY	1294	TATATTGCCAGAGCTGGATGACTCTCTGAGGAAATATGCTTGGTCCATCTCTACAT	1353	Db	721	LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTy	740
Db	361	TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyAlaTrpSerIleLeuLeuAsp	380	QY	2434	GGAGGATACCTCTCCCTGATGGCATTAATGACAGGTCAGATATCTTCAGGTTGCTATT	2493
QY	1354	CGTCCAGACTCCCTACAGATAGTGTGATCTCACTGCAATTTATTTATCCAGTAGAA	1413	Db	741	GlyGlyTyLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle	760
Db	381	ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu	400	QY	2494	GCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGATACACGGAACGTTATATG	2553
QY	1414	GATGATGTTATGGAAGGAGAGACTCATGAGTCAGTCAGTCGCTGATTCGAGCCACTA	1473	Db	761	AlaGlyAlaProValThrLeuTrpIlePheTyLysThrGlyThrGluArgTyMet	780
Db	401	AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu	420	QY	2554	GGTCACCTCGACCCAGAAATGAACAGGGCTATTACTTAGGATCTGTGGCATGCAAGCAGAA	2613
QY	1474	ATTATCTATGAAGAAACACAGACATCTGGATTAATATCCATGACATCTTTCATGTTT	1533	Db	781	GlyHisProAspGlnAsnGluGlnGlyTyLysLeuGlySerValAlaMetGlnAlaGlu	800
Db	421	IleIleTyroGluGluThrThrAspIleThrIleAsnIleHisAspIlePheHisValPhe	440	QY	2614	AAAGTTCCCTCTGAAACAAATCGTTTACTGCTCTTACATGTTTCTCTGGATGAGAATGTC	2673
QY	1534	CCCCAAGTCACGAAGAGAAATGAGTTTATTTTGGCTCTGATTCGAAACAGGTTTC	1593	Db	801	LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal	820
Db	441	ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe	460	QY	2674	CATTTTGCACATACCACTATATTTACTGAGTTTTTTAGTGAGGCTGGAAAGCCATATGAT	2733
QY	1594	CGTCATTTATACAAATTTACATCTATTTTAAAGGAAACAAATATAAOCATCAGTGGT	1653	Db	821	HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyAsp	840
Db	461	ArgHisLeuTyLysIleThrSerIleLeuLysGluSerLysTyLysArgSerSerGly	480	QY	2734	TTACAGATCTATCTCAGGAGAGACACAGATTAAGAGTTCTTGAATCGGAGAACATTTAT	2793
QY	1654	GGGTGCTCTCTCAAGTATTTCAAGTCTCTATCAAGAGGAGATAGCAATTTACCACT	1713	Db	841	LeuGlnIleTyroGlnGluArgHisSerIleArgValProGluSerGlyGluHisTy	860
Db	481	GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer	500	QY	2794	GAATCGCATCTTTTGACCTACTCTCAAGAAACCTTGGATCAGCTATTGCTGCTTAAAA	2853
QY	1714	GGTCAATGGGAAGTTCTTGGCGGATGATCTAATATCCAAAGTTGATGAAGTCAGAGG	1773	Db	861	GluLeuHisLeuLeuHisTyroGlnGluAsnLeuGlySerArgIleAlaAlaLeuLys	880
Db	501	GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg	520	QY	2854	GTGATA 2859	
QY	1774	CTGGTATTTTGAAGCCCAAGACTCCCTTTTAGAGCATCACCTGTACGTAGTCAGT	1833	Db	881	ValIle 882	
Db	521	LeuValTyroPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyroValValSer	540	RESULT 4			
QY	1834	TACGTAATCTCGAGAGTGACAGGCTGACCTGCGCTGCTACTCATCTTCTGCTGC	1893	AAE24170			
Db	541	TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyroSerHisSerCysCys	560	ID AAE24170	standard; protein; 882 AA.		
QY	1894	ATCAGTCAGCACTGTGACTCTTTTATAGTATAGTATAGTAAACCAGAAATCCACACTGT	1953	XX AAE24170;			
				AC			
				XX			

DT 23-SEP-2002 (first entry)
XX Human dipeptidyl peptidase 8 (DPP8) protein.
DE Human; dipeptidyl peptidase; DPP; neoplasia; type II diabetes; cirrhosis;
XX autoimmunity; human immuno deficiency virus; HIV infection; cytostatic;
KW graft rejection; antidiabetic; antiinflammatory; immunosuppressive;
KW antiviral; enzyme.
XX Homo sapiens.
OS WO200234900-A1.
PN 02-MAY-2002.
XX 29-OCT-2001; 2001WO-AU001388.
PF 27-OCT-2000; 2000AU-00001078.
PR (UNSY) UNIV SYDNEY.
XX Abbott CA, Gorrell MD;
PI WPI; 2002-454646/48.
DR N-PSDB; AAD38956.
XX New dipeptidyl peptidase (DPP) peptides, useful for screening inhibitors
PT of DPP catalytic activity, which may be employed to treat e.g. neoplasia,
PT type II diabetes, cirrhosis, autoimmunity, graft rejection and HIV
PT infection.
XX Example; Fig 1; 91pp; English.
XX The present invention relates to dipeptidyl peptidase (DPP) proteins and
CC polynucleotides encoding such proteins. The DPP peptides are useful for
CC screening inhibitors of DPP catalytic activity. The inhibitors are useful
CC for treating neoplasia, type II diabetes, cirrhosis, autoimmunity, graft
CC rejection and HIV (human immuno deficiency virus) infection. The present
CC sequence is human DPP8 protein
XX
XX Sequence 882 AA;
SQ
Alignment Scores:
Pred. No.: 0 Length: 882 -
Score: 4700.00 Matches: 882
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 84.7% Indels: 0
DB: 5 Gaps: 0
US-10-825-632-2 (1-3120) x AAE24170 (1-882)
QY 214 ATGGCAGCAGCAATGGAAACAGACAGCTGGGTGTGAGATATTGTAACCTGGGACTGT 273
DB 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluLeuPheGluThrAlaAspCys 20
QY 274 GAGGAGAAATGAATCACAGATCGGCCTAAATGGAGCCTTTTATGTTGAGCGGTAT 333
DB 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
QY 334 TCCTGGAGTCAGCTTAAAGCTGCTTGGCGATACAGAAATATCATGCTACATGATG 393
DB 41 SerTrpSerGlnLeuLysLeuLeuAlaPheThrArgLysTyrHisGlyTyrMetMet 60
QY 394 GCTAAGGCACCATGATTTTCATGTTGTGAAGAGAAATGATCCAGATGGACCTCATTC 453
DB 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
QY 454 GACAGATCTATTACCTGCCATGCTGTGTGAGAACAGAGAAATACACTGTTTATTCT 513
DB 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluLysAsnThrLeuPheTyrSer 100
QY 514 GAAATCCCAAAACTATCAATAGAGCAGCAGTCTTAATGCTCTCTTGGAGCCTCTTTTG 573

DB 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
QY 574 GATCTTTTTCAGCAACACTGGACTATGGAATGTATTCTCGAGAGAGAACTATTAAAGA 633
DB 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140
QY 634 GAAAGAAAACGATTTGGAAACAGTCGGAAATCTCTTACGATTTATCCCAAGGAGTGA 693
DB 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
QY 694 ACATTTCTCTTTCAAGCGGTAGTGAATTTATCACGTAAAGATGAGGCGCCACAGGA 753
DB 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
QY 754 TTTTACGCAACAACTTTTAAAGCCCAATCTAGTGGAAACTAGTTCTCCACATACGATG 813
DB 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
QY 814 GATCCAAAATTTATGCCCGCTGATCCAGACTGGATGCTTTTATACATACGACGATATT 873
DB 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
QY 874 TGGATATCTAATCATCTGAACACAGAGAGAAAGGAGACTCACTTATGTGCAATGAGCTA 933
DB 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu 240
QY 934 GCCAAATCGAAGAGATGCCAGATCAGCTGAGTGCCTACCTTTGTTCTTCAAGAGAA 993
DB 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu 260
QY 994 TTTGATAGATATTCGGCTATTGGTGTGTCCTCAAAAGCTGAAACACTCCCGAGTGGT 1053
DB 261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280
QY 1054 AAATTTCTAGAAATCTATATCAAGAAATGATGATCTGAGGTGGAAATTTATCATGTT 1113
DB 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300
QY 1114 ACATCCCTATGTTGGAAACAAGGAGGCGAGATTCATTTCCGTATCTCTTAAACAGGTACA 1173
DB 301 ThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThr 320
QY 1174 GCAAAATCCTAAAGTCACTTTTAAGATGTCAGAAATAATGATGATGCTGAAGGAGATC 1233
DB 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
QY 1234 ATAGATGTCATAGATAAGGAACTAATTCAACTTTTGGAGATTCATTTTGAAGGAGTGA 1293
DB 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
QY 1294 TATATTCCCAAGCTGGATGGACTCCTGAGGGAAATAATGCTTGTGCTCATCTACTAGAT 1353
DB 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
QY 1354 CGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATTTATTTATCCAGTAGAA 1413
DB 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
QY 1414 GATCATGTTATGGAAGGCGAGAGACTCATTGAGTCAGTGCCTGATTCTGTGACGCGACTA 1473
DB 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
QY 1474 ATTATCTATCAAGAAACAACAGACATCTGGATAAATAATCCATGACATCTTTTCATGTTT 1533
DB 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
QY 1534 CCCCAGAGTCACGAGAGGAAATGAGTTATTTTGGCTCTGAATGCAAAACAGCTTTC 1593
DB 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
QY 1594 CGTCATTTATCAAAATTTACATCTATTTTAAAGGAGCAAAATATAAAGATCCAGTGTGT 1653

Db 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
QY 1654 GGGCTGCTCTCCAAAGTGAATTCCAAGTGTCTTCAAGAGAGAGATAGCAATTCACAGT 1713
Db 481 GlyLeuProAlaProSerAspPheLysCysProIleGlyGluIleAlaIleThrSer 500
QY 1714 GGTGAATGGGAAGTTCTTTGGCCGCATCGATCTAATATCAAGTTGATCAAGTCAGAAAG 1773
Db 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
QY 1774 CTGGTATATTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTACGTAGTCACT 1833
Db 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSer 540
QY 1834 TACGTAAATCTCGAGAGCTCACAAGCTGACTGACCTGGCTACTACATTCCTTCTGCTGC 1893
Db 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
QY 1894 ATCAGTCCAGCACTGTGACTTCTTTATAGTAAGTATAGTAACCAAGAGATCCACACTGT 1953
Db 561 IleSerGlnHisCysAspPheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
QY 1954 GTGCTCCTTACAACTATCAAGTCCCTGAGAGTACCACCTGCAAAACAAGCAATTT 2013
Db 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
QY 2014 TGGGCACCAATTTGGATTACAGAGTCTCTCTCTGACTATATCTCTCCAGAAATTTTC 2073
Db 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 620
QY 2074 TCTTTTGAAGTACTCTGGAATTTACATTTGATGGGATGCTCTCAAGCTCATGATCTA 2133
Db 621 SerPheGluSerThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
QY 2134 CAGCTTGGAAAGAAATATCTCTACTGTCTGTTCATATATATGTTGGTCTCAGGTGCAGTTG 2193
Db 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 660
QY 2194 GTGAATAATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCCCTAGCCTCTCTAGT 2253
Db 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680
QY 2254 TATGTGTTGTAGTCATAGACACAGGGGATCCTGTCCAGGGGCTTAAATTTGAAGC 2313
Db 681 TyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 700
QY 2314 GCCTTTAATATATAATGGTCAATAGAAATGACGATCAGGTGGAAGACTCCCAATAT 2373
Db 701 AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr 720
QY 2374 CTAGCTTCTCGATATGATTTCAATTCAGTTCGCTGCGGCATCCACGCTGGCTTAT 2433
Db 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyIleTrpSerTyr 740
QY 2434 GGAGGATACCTCTCCCTGATGGCATTAATGACAGAGTCAGATATCTTCAGGGTTCGTATT 2493
Db 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
QY 2494 GCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGATACAGGAAGCTTATATG 2553
Db 761 AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780
QY 2554 GGTCACTCCCTGACAGAAAGACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAA 2613
Db 781 GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
QY 2614 AAGTTCTCCCTCTGAACCAAAATCGTTTACTGCTCTTACATGTTTCTCTGATGAGATGTC 2673
Db 801 LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
QY 2674 CATTTTGACATACAGATATATTACTGAGTTTTTTAGTGGGGCTGGAAGCCATATGAT 2733
Db 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840

QY 2734 TTACAGATCTATCTCTCAGGAGACACAGACGATTAAGAGTTCTCTGAATCGGAGACATTAT 2793
Db 841 LeuGlnIleTyrProGlnGluArgHiserIleArgValProGluSerGlyGluHisTyr 860
QY 2794 GAAGTCGATCTTTTGCATCTACCTTCAAGAAAACCTTGGATCACGATATTCGCTCTAAAA 2853
Db 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLys 880
QY 2854 GTGATA 2859
Db 881 ValIle 882
RESULT 5
AAU74749
ID AAU74749 standard; protein; 882 AA.
XX AAU74749;
AC AAU74749;
XX
DT 09-APR-2002 (first entry)
XX
DE Human protease PRTS-9 protein sequence.
XX
KW Human; protease; PRTS; gastrointestinal; Crohn's disease; cancer;
KW cardiovascular; atherosclerosis; autoimmune disorder; dermatitis;
KW inflammatory disorder; acquired immunodeficiency syndrome; AIDS;
KW cell proliferative disorder; developmental disorder; epilepsy;
KW Duchenne muscular dystrophy; epithelial disorder; neurological disorder;
KW reproductive disorder; endometriosis.
XX
OS Homo sapiens.
XX
PN WO200198468-A2.
XX
PD 27-DEC-2001.
XX
PF 13-JUN-2001; 2001WO-US019178.
XX
PR 16-JUN-2000; 2000US-0212336P.
PR 22-JUN-2000; 2000US-0213953P.
PR 29-JUN-2000; 2000US-0215396P.
PR 07-JUL-2000; 2000US-0216821P.
PR 14-JUL-2000; 2000US-0218946P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Yue H, Elliott VS, Gandhi AR, Lal P, Au-Young J, Tribouley CM;
PI Deleage AM, Baughn MR, Nguyen DB, Lee EA, Hafalia A, Khan PA;
PI Walia NK, Yao MG, Lu DAM, Patterson C, Tang YT, Walsh RT;
PI Azimzai Y, Lu Y, Ramkumar J, Xu Y, Reddy R, Das D, Kearney L;
PI Kallick DA;
XX
DR WPI; 2002-090437/12.
DR N-PSDB; ABK12892.
XX
XX Twenty one human proteases (referred to as PRTS-1 to PRTS-21), useful in
XX the diagnosis, treatment and prevention of gastrointestinal (e.g.
XX gastritis), cardiovascular (e.g. atherosclerosis) and cell proliferative
XX (e.g. cancer) disorders.
XX
XX Claim 1; Page 140-142; 17pp; English.
XX
XX The present invention relates to twenty one new human proteases, referred
XX to as PRTS-1 to PRTS-21. The PRTS polynucleotides and polypeptides of the
XX invention are useful in the diagnosis, treatment and prevention of
XX gastrointestinal e.g. gastritis, esophageal carcinoma and Crohn's
XX disease, cardiovascular e.g. atherosclerosis, hypertension and myocardial
XX infarction, autoimmune/inflammatory e.g. acquired immunodeficiency
XX syndrome (AIDS), allergies and osteoarthritis, cell proliferative e.g.
XX cancer, developmental e.g. Duchenne and Becker muscular dystrophy,
XX epithelial e.g. dermatitis, neurological e.g. epilepsy and Alzheimer's
XX disease and reproductive e.g. infertility and endometriosis disorders.
XX Numerous other examples of each disorder are given in the specification.

CC The present protein sequence represents the human protease PRTS-9 protein
CC of the invention
XX
SQ Sequence 882 AA;

Alignment Scores:
Pred. No.: 0 Length: 882
Score: 4700.00 Matches: 882
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 84.7% Indels: 0
Dbs: 5 Gaps: 0

US-10-825-632-2 (1-3120) x AAU74749 (1-882)

QY	214	ATGCGAGCAGCAATGGAACAGACAGCTGGGTGTTGAGATATTGAACTCGGACTGT	273
DB	1	MetAlaAlaAlaMetGluThrGluInLeuGlyValGluIlePheGluThrAlaAspCys	20
QY	274	GAGGAGAAATATTGAATCACAGGATCGGCCCTAAATTTGGAGCCTTTTATGTCAGCGGTAT	333
DB	21	GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr	40
QY	334	TCCTGGATCAGCTTAAAGCTCTTGGCCGATACAGAAATATCATCGCTACATGATG	393
DB	41	SerTrpSerGlnLeuLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet	60
QY	394	GCTAAGGCACCATGATTTTCATGTTGTGAAGAGGAATGATCCAGATGGACTCATTTCA	453
DB	61	AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer	80
QY	454	GACAGAATCTATTACCTTGCATGCTGGTGAGAACAGAGAAATACATGTTTATTCT	513
DB	81	AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeupheTyrSer	100
QY	514	GAATTTCCAAATCATCATAGAGCAGCAGCTTATGCTCTCTTGGAGCCTCTTTTG	573
DB	101	GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu	120
QY	574	GATCTTTTCAGGCAACACTGACTATGGAATGATTCTCGAGAAGAAGAACTATTAA	633
DB	121	AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg	140
QY	634	GAAGAAACGANTTGGACAGTCGGAATTCCTTTACGATTATCACCAAGAAATGA	693
DB	141	GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly	160
QY	694	ACATTTCTGTTTCAAGCCGTTAGTGAATTTATCACGTAAAGATGGAGGCCACAGGA	753
DB	161	ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly	180
QY	754	TTTACGCAACACCTTTAAGCCCAATCTAGTGGAACTAGTGTCTCCCAACATACGATG	813
DB	181	PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet	200
QY	814	GATCCAAATATTATGCCCGCTGATCCGACTGGATTGCTTTTATACATACCAACGATAT	873
DB	201	AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle	220
QY	874	TGATATCTAACATCTGAACAGAGAAAGGAGACTCATCTTATGTGCAATGAGCTA	933
DB	221	TrpIleSerAsnIleValThrArgGluGluArgArgLeuThrTyrValHisAsnGluLeu	240
QY	934	GCCAACTGAAGAAGATGCCGATCAGCTGAGTCGCTACCTTTGTTCTCCAGAGAA	993
DB	241	AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaIleThrPheValLeuGlnGlu	260
QY	994	TTTGTAGATATTCTGGCTATTGGTGTCTCCAAAGCTGAAACAACTCCCGAGTGGT	1053
DB	261	PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly	280
QY	1054	AAAATTCTTAGAATCTTATATGAAGAAATGATGAATCTGAGGTGGAAATTTATTCAT	1113

DB	281	LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal	300
QY	1114	ACATCCCTATGTTGGAAACAAAGAGGGCAGATTTCATTCGTTATCCTAAACAGGTACA	1173
DB	301	ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr	320
QY	1174	GCAATCCTAAAGTCACCTTTTAAGATGTCAGAAATATGATTGATGCTGAGGAAGATC	1233
DB	321	AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle	340
QY	1234	ATAGATGTCATAGATTAAGAACTAATTCACCTTTTTCAGATTTCTATTGTAAGGAGTTGAA	1293
DB	341	IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu	360
QY	1294	TATATTGCCAGAGTCGATGACTCTCTGAGGAAATATGCTGGTCCATCTCTACTAGAT	1353
DB	361	TyrIleAlaArgAlaGlyTyrTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp	380
QY	1354	CGCTCCAGACTCGCCTACAGATAGTGTGATCTACCTGAAATATTATTTATCCAGTAGAA	1413
DB	381	ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu	400
QY	1414	GATGATGTTATCGAAAGCAGAGACTCATTTGAGTCAGTCAGTCGCTGATCTGTCGCGCACTA	1473
DB	401	AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu	420
QY	1474	ATTATCTATGAAGAACACAGACATCTGGATAAATATCCATGACATCTTTCTATGTTTTT	1533
DB	421	IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe	440
QY	1534	CCCAAGTCACGAAGAGGAATTTAGTTTATTTTGGCTCTGAATGCAAAACAGGTTTC	1593
DB	441	ProGlnSerHisGluGluGluIleGluPheIleAspSerGluCysLysThrGlyPhe	460
QY	1594	CGTCATTTATACAAATATCATCTATTTTAAAGGAACAAATAAATACCATCCAGTGGT	1653
DB	461	ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly	480
QY	1654	GGCTGCTGCTCCCAAGTATTTCAAGTGCTCTATCAAGAGGAGATAGACATTAACCAT	1713
DB	481	GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer	500
QY	1714	GGTGAATGGGAAGTCTTTCGCGCATGATCTAATATCCAAGTTGATGAAGTCAGAGG	1773
DB	501	GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg	520
QY	1774	CTGGTATTTTGAAGGCCACCAAGACTCCCTTTAGAGCATCACCTGTAGTAGTCAGT	1833
DB	521	LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSer	540
QY	1834	TACGTAAATCCTGAGAGGTGACAGGCTGACTGACCTGGCTACTCATCTTCTGCTGC	1893
DB	541	TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys	560
QY	1894	ATCAGTCAGACACTGCTCTTTTATAAGTAAGTATAGTAACCAAGAAATCCACACTGT	1953
DB	561	IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys	580
QY	1954	GTGTCCCTTTTACAGCTATCAAGTCTCCTGAAGATGACCCCACTTGCAGAAACAAAGAA	2013
DB	581	ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe	600
QY	2014	TGGGCCACCAATTTTGGATTTCAGAGGTCTCTCTGCTGACTATCTCTCCAGAAATTTTC	2073
DB	601	TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePhe	620
QY	2074	TCCTTTCAAGAGTACTCTGGATTTACATTTGATATGGGATGCTCTACAAGCCTCATGATCTA	2133
DB	621	SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu	640
QY	2134	CAGCTCGAAAGAAATATCTACTGTGCTGTTTCATATATGTTGGTGGTCTCAGGTGAGTTG	2193
DB	641	GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu	660

QY 2194 GTGAATATCGGTTTAAAGAGTCAAGTATTTCGCTTGGAATACCTAGCCTCTTAGGT 2253
Db |||||||
QY 2254 TATGCTGTTGATGATAGACACACAGGGGATCCTGTCCACGAGGGCTTAAATTTGAGGC 2313
Db |||||||
QY 2314 GCCTTTAAATATAAATGGGTCAAATAGAAATGACGATCAGGTGGAAGGACTCCCAATAT 2373
Db |||||||
QY 2374 CTAGCTTCTCGATATGATTTCATTGACTTGTGATCGTGTGGCCATCCACGGCTGGTCCATAT 2433
Db |||||||
QY 2434 GGAGGATACCTCTCCCTGATGGCATTAAATGACAGGGTCAGATATCTTCAGGGTTGCTATT 2493
Db |||||||
QY 2494 GCTGGGGCCCACTCTCTGAGTCTTCTATGATACAGGATACACGGAAAGCTTATATG 2553
Db |||||||
QY 2554 GGTACACCTGACAGAGTGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAA 2613
Db |||||||
QY 2614 AAGTTCCCTCTGACCAAAATCGTTTACTGCTCTTACATGGTTTCTCGATCAGATGTC 2673
Db |||||||
QY 2674 CATTTGACATACAGTATATTAATGATGTTTCTAGTGGCTGGAAGCCATATGAT 2733
Db |||||||
QY 2734 TTACAGATCTATCTCTAGAGAGACACAGCATTAAGATTCCTGAATCGGAGAACTATAT 2793
Db |||||||
QY 2794 GAACTGCATCTTTGCACTACCTTCAAGAAAACCTTGGATCAGCTATCTCTCTTAAAA 2853
Db |||||||
QY 2854 GTGATA 2859
Db |||||||
RESULT 6
ID ADI17086
AC ADI17086; standard; protein; 882 AA.
DT 15-APR-2004 (first entry)
DE Human NOVX protein homologue SeqID 622.
KW human; NOVX; cardiomyopathy; atherosclerosis; cancer; diabetes;
KW inflammation; autoimmunity disorder; allergy; blood disorder;
KW acquired immunodeficiency syndrome; AIDS; obesity; asthma;
KW immunoglobulin (Ig)A nephropathy; cirrhosis; arthritis;
KW Alzheimer's disease; infection; str.
OS Homo sapiens.
FN WO200268649-A2.
PD 06-SEP-2002.
PF 31-JAN-2002; 2002WO-US002785.
XX

PR 31-JAN-2001; 2001US-0265395P.
PR 31-JAN-2001; 2001US-0265412P.
PR 31-JAN-2001; 2001US-0265514P.
PR 02-FEB-2001; 2001US-0266406P.
PR 05-FEB-2001; 2001US-0266676P.
PR 07-FEB-2001; 2001US-0266975P.
PR 08-FEB-2001; 2001US-0267057P.
PR 09-FEB-2001; 2001US-0267823P.
PR 15-FEB-2001; 2001US-0268974P.
PR 26-FEB-2001; 2001US-0271664P.
PR 27-FEB-2001; 2001US-0271839P.
PR 27-FEB-2001; 2001US-0271855P.
PR 02-MAR-2001; 2001US-0272788P.
PR 14-MAR-2001; 2001US-0273046P.
PR 14-MAR-2001; 2001US-0275925P.
PR 14-MAR-2001; 2001US-0275947P.
PR 14-MAR-2001; 2001US-0275950P.
PR 15-MAR-2001; 2001US-0275989P.
PR 15-MAR-2001; 2001US-0276448P.
PR 15-MAR-2001; 2001US-0276450P.
PR 16-MAR-2001; 2001US-0276397P.
PR 16-MAR-2001; 2001US-0276768P.
PR 20-MAR-2001; 2001US-0278652P.
PR 26-MAR-2001; 2001US-0278775P.
PR 29-MAR-2001; 2001US-0279882P.
PR 29-MAR-2001; 2001US-0279884P.
PR 30-MAR-2001; 2001US-0280147P.
PR 11-APR-2001; 2001US-0282392P.
PR 11-APR-2001; 2001US-0283083P.
PR 23-APR-2001; 2001US-0285133P.
PR 03-MAY-2001; 2001US-0285743P.
PR 03-MAY-2001; 2001US-0288327P.
PR 29-MAY-2001; 2001US-0288504P.
PR 30-MAY-2001; 2001US-0294047P.
PR 08-JUN-2001; 2001US-0296964P.
PR 18-JUN-2001; 2001US-0298959P.
PR 19-JUN-2001; 2001US-0299324P.
PR 13-AUG-2001; 2001US-0312020P.
PR 16-AUG-2001; 2001US-0312889P.
PR 16-AUG-2001; 2001US-0312908P.
PR 21-AUG-2001; 2001US-0313390P.
PR 28-AUG-2001; 2001US-0315470P.
PR 31-AUG-2001; 2001US-0316447P.
PR 07-SEP-2001; 2001US-0318115P.
PR 12-SEP-2001; 2001US-0318118P.
PR 19-SEP-2001; 2001US-0318740P.
PR 18-OCT-2001; 2001US-0323379P.
PR 18-OCT-2001; 2001US-0330245P.
PR 14-NOV-2001; 2001US-0330308P.
PR 14-NOV-2001; 2001US-0332701P.
PA (CURA-) CURAGEN CORP.
XX Tchernev VT, Spytek KA, Zerhusen BD, Patturajan M, Shinkets RA;
XX Li, Ganglioni KA, Padigaru M, Anderson DM, Rastelli L, Miller CE;
XX Gerlach VL, Taupier RJ, Gusev VY, Colman SD, Wolenc AR, Pena CEA;
XX Furtak K, Grosse RW, Alsobrook JF, Lepley DM, Rieger DK, Burgess CE;
XX WPI; 2002-706998/76.
XX
XX New NOVX polypeptides and nucleic acids, useful for preventing or
XX treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,
XX atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or
XX pharmacogenomics.
XX
XX Disclosure; SEQ ID NO 622; 1498pp; English.
XX
XX This invention relates to a novel nucleic acids, and encoded polypeptides
XX thereof, which have properties related to the stimulation of biochemical
XX

or physiological responses in a cell, tissue, organ or organism.
Specifically, it refers to the use of biologically active fragments for
diagnostic and prognostic assays and furthermore in the treatment of
diverse pathological conditions. The present invention describes novel
human and murine NOVX proteins, as well as methods to modulate their
expression using antisense oligos, ribozymes and peptide nucleic acids.
The NOVX polypeptides, polynucleotides and antibodies are useful in
treating or preventing NOVX-associated disorders, e.g. cardiomyopathy,
atherosclerosis, cancer and diabetes. Furthermore, they may be used in
treating or preventing diseases such as inflammation, autoimmune
disorders, allergies, blood disorders, acquired immunodeficiency syndrome
(AIDS), obesity, asthma, immunoglobulin (IgA) nephropathy, cirrhosis,
arthritis, Alzheimer's disease, infections, stroke, muscular dystrophy
and epilepsy. Accordingly, these molecules have many activities including
cytostatic, cardiant, antiinflammatory, immunosuppressive, antiallergic,
haemostatic, anti-HIV, antidiabetic, antiarteriosclerotic, anorectic,
antiasthmatic, nephrotropic, antiarthritic, hepatotropic,
neuroprotective, nontropic, antibacterial, virucide, antiparasitic,
relaxant and anticonvulsant. In addition, they are useful in screening
assays to identify small molecules that modulate or inhibit, for example,
neurogenesis, wound healing and angiogenesis. The nucleic acids are also
used as in chromosome mapping, tissue typing, preventive medicine and
pharmacogenomics. This polypeptide is a homologue of a human NOVX protein
of the invention.

Sequence 882 AA;

Alignment Scores:

Pred. No.:	0	Length:	882
Score:	4700.00	Matches:	882
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	84.7%	Indels:	0
DB:	5	Gaps:	0

US-10-825-632-2 (1-3120) x ADI17086 (1-882)

QY	214	ATGCGACGACCAATGGAAACAGAACAGCTGGGTGGTGGATATTTGAACTGCGGACTGT	273
DB	1	MetAlaAlaMetGluThrGluGlnLeuGlyValGluLeuPheGluThrAlaAspCys	20
QY	274	GAGGAGATATTCGAATCACAGGATCGGCTAAATGGAGCCCTTTTATGTTGACGGTAT	333
DB	21	GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr	40
QY	334	TCCTGGAGTCAGCTTAAAAAGCTGCTTGGCATCCAGAAATATCATGGCTACATCATG	393
DB	41	SerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet	60
QY	394	GCTAAGGCACCATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTTCA	453
DB	61	AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer	80
QY	454	GACAGAATCTATTACCTTGGCATGTCTGGTGAGAACAGAGAAATATACATCTTTTATCT	513
DB	81	AspargileTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer	100
QY	514	GAATTCCTCCAAAATATCAATAGACGAGCTCTTAATGCTCTCTTGGAGCCCTCTTTG	573
DB	101	GluileProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu	120
QY	574	GATCTTTTTCAGGCACACTGGACTATGGAATGCTTCTTACGATTATCACCAAGGAGTGA	633
DB	121	AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg	140
QY	634	GAAGAAACCGCATTTGGAACAGTCGGAATGCTTCTTACGATTATCACCAAGGAGTGA	693
DB	141	GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly	160
QY	694	ACATTTCTGTTTCAAGCCGGTAGTGGAATTTATACGTAAAGATGAGGGCCCAAGGA	753
DB	161	ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly	180

QY	754	TTTAGCAACAACTTAAAGGCCCAATCTAGTGGAACTAGTTGTGCCCAACATAGCGATG	813
DB	181	PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet	200
QY	814	GATCCAAAATATATGCCCGCTGATCCAGACTCGATTGCTTTTATACATAGCAACATATT	873
DB	201	AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle	220
QY	874	TGGATATCTTAACATCTGTAACAGAGAAAGAGACATCATTTATGTGCACAATAGACTA	933
DB	221	TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu	240
QY	934	GCCAACTGGAAGAAGATGCCAGATCAGCTGAGCGCTGCTTCTTCTCCAGAGAA	993
DB	241	AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu	260
QY	994	TTTCATAGATATTCTGCTGCTATTCGCTCCAAAGCTGAAACAACTCCCAAGTGGTGGT	1053
DB	261	PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly	280
QY	1054	AAAAATCTTAGAATTTCTATATGAAGAAATGATGAATCTGAGTGGAAATTTATCATGTT	1113
DB	281	LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal	300
QY	1114	ACATCCCTATGTTGGAACAGAGGCGCAGATTCAATCCGTTATCTTAAACAGGTACA	1173
DB	301	ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr	320
QY	1174	GCAATCCTAAAGTCATCTTTTAAGATGTCAGAAATAATGATTGATGCTGAAGGAGGATC	1233
DB	321	AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle	340
QY	1234	ATAGATGCTCATAGATAAGGAACCTAAATCAACCTTTTGGAGATTTCTATTTGAAGGAGTGA	1293
DB	341	IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu	360
QY	1294	TATATGCCAGAGCTGAGTGGACTCTGAGGGGAAATAATATGCTGGTCCATCTCTACTAGAT	1353
DB	361	TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp	380
QY	1354	CGCTCCAGAGCTCGCTACAGATAGTGTGATCTCACCTGGAATTTATTTATCCAGTAGAA	1413
DB	381	ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu	400
QY	1414	GATCATGTTATGGAAGCGACAGACTCATTTGAGTCCAGTCCCTGATTTCTGTGACGCCACTA	1473
DB	401	AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu	420
QY	1474	ATTATCTATGAAGAAACACAGACATCTCGATAAATATCCATGACATCTTTCATGTTTTT	1533
DB	421	IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe	440
QY	1534	CCCCAAGTCACGAGAGGAATTTGATTTATTTTTCCTCTGATGCAATGCAAAACAGTTTC	1593
DB	441	ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe	460
QY	1594	CGTCATTTATACAAAATTTACATCTATTTTAAAGGAAGCAAAATATAAAGCATCCAGTGGT	1653
DB	461	ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly	480
QY	1654	GGGCTGGCTGCTCCAACTGATTTTCAAGTGTCTTATCAAAAGAGAGATAGCAATACCACT	1713
DB	481	GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer	500
QY	1714	GGTGATGGGAAGTTCTTGGCCGCATGATCTAATATATCCAAAGTTGATGAGTCAAGG	1773
DB	501	GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg	520
QY	1774	CTGTGTATTTTGAAGSCCAAGACTCCCTTTTAGAGCATCACCTGTACCTGTAGTCACT	1833
DB	521	LeuValTyrPheGluGluThrLysAspSerProLeuGluHisLeuTyrValValSer	540
QY	1834	TACGTAAATCTCGAGAGGTGCAAGGCTGACTGACCGTGGCTACTCACAATCTTCTGCTGC	1893

Db	541	TyrValAanProGlyGluValThrArgLeuThrAspArgGlyTyr-SerHisSerCysCys	560	XX	ABU07720 standard; protein; 882 AA.
Qy	1894	ATCAGTCAGCAGCTGCTCTTTATAGTAAGTAGTAACAGAGAAGTAATCCACACTGT	1953	AC	ABU07720;
Db	561	IleSerGlnHisCysAspPheIleSerLysTyr-SerAenGlnLysAenProHisCys	580	XX	19-MAY-2003 (first entry)
Qy	1954	GTGTCCTCCCTTACAGCTATCAAGTCCTGAGATGAGCCCACTTGCAAAACAAGAAATTT	2013	DE	Human serine protease HIPHUM46.
Db	581	ValSerLeuLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe	600	XX	Human; enzyme; HIPHUM46; serine protease; gene therapy; osteoarthritis;
Qy	2014	TGGGCACATTTGGATTGAGAGCTCTTCCCTGACTATATCTCCACAAATTTTC	2073	KW	serine protease activity modulation; dipeptidyl peptidase activity;
Db	601	TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspLysThrProGluLlePhe	620	KW	musculoskeletal disease; Hepatitis B virus infection; myotonic dystrophy;
Qy	2074	TCTTTTGAAGPACTCTGGATTAATGTTATGGGATGCTCTACAGGCTCATGATCTA	2133	KW	Alzheimer's disease; parasupranuclear palsy; Huntington's disease;
Db	621	SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAapLeu	640	KW	amyotrophic lateral sclerosis; malabsorption syndrome; lung disease;
Qy	2134	CAGCCTGGAAGAAATATCTACTGCTGCTTCATATATGATGGTGGTCTCAGGTGCAGTTG	2193	KW	irritable bowel syndrome; type I diabetes; faecal incontinence; anaemia;
Db	641	GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu	660	KW	haemorrhoid; proctitis; rectal polyp; small bowel tumour; dyslexia;
Qy	2194	GTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCCCTAGCCTCTCAGGT	2253	KW	colorectal tumour; ceroid lipofuscinosis; allergic encephalomyelitis;
Db	661	ValAenAenAenAenAenAenAenAenAenAenAenAenAenAenAenAenAenAenAen	680	XX	multiple sclerosis.
Qy	2254	TATGTGTTGTAGTATAGACACACAGGGGATCCTGTCCAGGAGGCTTAAATTTGAAGC	2313	OS	Homo sapiens.
Db	681	TyrValValValValValValValValValValValValValValValValValValVal	700	XX	Key
Qy	2314	GCCTTTAATATAATGGTCAATAGAAATTCACGATCAGGTGGAAGACTCCCAATAT	2373	PH	Location/Qualifiers
Db	701	AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr	720	FT	259..260
Qy	2374	CTAGCTCTCGATATGATTCATTCAGCTAGATCGTGGGCGATCCACGGCTGCTCTAT	2433	FT	Note= "Paired glutamates of the beta propeller domain"
Db	721	LeuAlaSerArgTyrAspPheIleAspLysAspArgValGlyIleHisGlyTyrSerTyr	740	FT	739
Qy	2434	GGAGGATACCTCTCCCTGATGCGATTAATGACAGAGTCAGATATCTCAGGGTTGCTATT	2493	FT	/label= Catalytic_serine_residue
Db	741	GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheAGValAlaIle	760	FT	617
Qy	2494	GCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGATACAGGACCTTATATG	2553	FT	/label= Catalytic_aspartate_residue
Db	761	AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet	780	FT	849
Qy	2554	GGTCACCCCTGACACAGAAATCAACAGGGCTATTACTAGGATCTGTGGCCATGCAAGCAGAA	2613	FT	/label= Catalytic_histidine_residue
Db	781	GlyHisProAspGlnAenGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu	800	XX	GB2374869-A.
Qy	2614	AAAGTTCCCTCTCAACCAATCGTTTACTGCTCTTACATGTTTCTCTGGATGAAATGTC	2673	XX	30-OCT-2002.
Db	801	LysPheProSerGluProAenArgLeuLeuLeuLeuHisGlyPheLeuAspGluAenVal	820	CC	22-JAN-2002; 2002GB-00001404.
Qy	2674	CATTTTGCACATACCATGATATTTACTGAGTCTTTTGTAGGGGCTGGAAAGCCATATGAT	2733	CC	23-JAN-2001; 2001GB-00001760.
Db	821	HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp	840	CC	(GLAX) GLAXO GROUP LTD.
Qy	2734	TTACAGATCTATCTCAGGAGACACAGCATTAAGATTCTGTAATCGGAGAACATTAT	2793	CC	Edbrooke MR, Lewis AP;
Db	841	LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr	860	CC	WPI; 2003-150703/15.
Qy	2794	GAATGATCTTTTGCACCTCTCAAGAAACCTTGGATCAGCTATTGCTGCTCTAAAA	2853	CC	N-PSDB; ABX12255.
Db	861	GluLeuHisLeuLeuHisTyrLeuGlnGluAenLeuGlySerArgIleAlaLeuLys	880	CC	Identifying modulators of serine protease activity useful for treating
Qy	2854	GTGATA 2859		CC	musculoskeletal diseases, by contacting cell expressing a novel serine
Db	881	ValIle 882		CC	protease polypeptide with a compound and monitoring serine protease
RESULT 7				XX	activity.
ABU07720				PS	Claim 10; Page 26-29; 30pp; English.
				XX	The invention relates to a method of identifying a substance that
				CC	modulates serine protease activity, comprising contacting a cell such as
				CC	a neuronal cell, lung cell, intestinal cell or a cell infected with a
				CC	virus, expressing a serine protease polypeptide (HIPHUM 46), or its
				CC	variant having dipeptidyl peptidase activity, or a serine protease
				CC	isolated from the cell with a test substance, and monitoring for serine
				CC	protease activity. The method is useful for identifying a substance that
				CC	modulates serine protease activity. A modulator of the serine protease is
				CC	useful in the manufacture of a medicament for treatment of fibrosia of
				CC	a musculoskeletal disease e.g. osteoarthritis, Hepatitis B virus
				CC	infection, Alzheimer's disease, parasupranuclear palsy, myotonic
				CC	dystrophy, Huntington's disease or amyotrophic lateral sclerosis.
				CC	Additional disease that may be treated using modulators of the serine
				CC	protease include malabsorption syndromes, irritable bowel syndrome, lung
				CC	disease, type I diabetes, faecal incontinence, haemorrhoids, proctitis,
				CC	rectal polyps, small bowel tumours, colorectal tumours, anaemia,
				CC	dyslexia, ceroid lipofuscinosis, allergic encephalomyelitis, and multiple
				CC	sclerosis. The present sequence represents the amino acid sequence of the
				CC	human serine protease HIPHUM46
				XX	Sequence 882 AA;
				SQ	

CC human and murine NOVX proteins, as well as methods to modulate their
 CC expression using antisense oligos, ribozymes and peptide nucleic acids.
 CC The NOVX polypeptides, polynucleotides and antibodies are useful in
 CC treating or preventing NOVX-associated disorders, e.g. cardiomyopathy,
 CC atherosclerosis, cancer and diabetes. Furthermore, they may be used in
 CC treating or preventing diseases such as inflammation, autoimmune
 CC disorders, allergies, blood disorders, acquired immunodeficiency syndrome
 CC (AIDS), obesity, asthma, immunoglobulin (Ig)A nephropathy, cirrhosis,
 CC arthritis, Alzheimer's disease, infections, stroke, muscular dystrophy
 CC and epilepsy. Accordingly, these molecules have many activities including
 CC cytostatic, anti-HIV, antidiabetic, antidiarrhoeal, antiallergic,
 CC haemostatic, nephrotropic, antithrombotic, antihypertensive, anorectic,
 CC antiasthmatic, neurotropic, antitubercular, virucide, antiparasitic,
 CC neuroprotective, nootropic, antibacterial, in addition, they are useful in screening
 CC assays to identify small molecules that modulate or inhibit, for example,
 CC neurogenesis, wound healing and angiogenesis. The nucleic acids are also
 CC used as in chromosome mapping, tissue typing, preventive medicine and
 CC pharmacogenomics. This polypeptide is a homologue of a human NOVX protein
 CC of the invention.

XX
 SQ Sequence 883 AA;

Alignment Scores:
 Pred. No.: 0 Length: 883
 Score: 4528.50 Matches: 845
 Percent Similarity: 97.7% Conservations: 18
 Best Local Similarity: 95.7% Mismatches: 19
 Query Match: 81.6% Indels: 1
 DB: 5 Gaps: 1

US-10-825-632-2 (1-3120) x ADI117085 (1-883)

QY 214 ATGGCAGCAGCAATGGAACAGAACAGCTGGTGTTCAGATATTTGAACCTGGGACTGT 273
 DB 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluLeuPheGluThrAlaGluCys 20
 QY 274 GAGGAG---ATATTGATCATCAGATCGGCTTAATGGAGCTTTTATGTGAGCGG 330
 DB 21 GluGluGlyAsnGlyGluSerGlnAspArgProLysLeuGluProPheThrValGluAsp 40
 QY 331 TATTCCTGGAGTCAGCTAAAGAGCTGCTCCGATACCAAGAAATATCATGCTACATG 390
 DB 41 TyrSerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMet 60
 QY 391 ATGGCTAAGCCACCATGATTTCATGTTTGTGAAGAGGAATGATCAGATGACCTCAT 450
 DB 61 MetAlaLysAlaProHisAspPheMetPheValLysArgThrAspProAspArgProHis 80
 QY 451 TCAGACAGAACTATTACCTGCTGATGCTGGTGAGACAGAGAAATACACTGTTTAT 510
 DB 81 SerAspArgValTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyr 100
 QY 511 TCTGAATTTCCAAACTATCAATAGCAGCAGCTTAAATGCTCTCTTGAAGCTCTT 570
 DB 101 SerGluLeuProLysThrIleAsnArgAlaValLeuMetLeuSerTrpLysProLeu 120
 QY 571 TTGGATCTTTTTCAGCAACACTGGACTATGGAATGTATTCTCGAAGAGAACTATTAT 630
 DB 121 LeuAspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeu 140
 QY 631 AGAAGAAGAAAGCGATTGGAACTGCGAATTCCTTACGATTATCCAAAGGAAGT 690
 DB 141 ArgGluArgLysArgIleGlyThrValGlyIleAlaLafyrAspTyrHisProGlySer 160
 QY 691 GGAACATTTCTGTTTCAACCGGTAGTGAATTTATCAGTAAAGATGGGGCCACAA 750
 DB 161 GlyThrPheLeuPheGlnAlaGlySerGlyIleTyrHisIleLysAspGlyGlyProHis 180
 QY 751 GGAATTTACCAACACTTTAAGGCCAATCTAGTGGAAACTAGTGTGCCAACATACGG 810
 DB 181 GlyPheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArg 200

QY 811 ATGGATCCAAATTTATGCCCCCGCTGATCCAGACTGGATTGCTTTTATACATACCAACGAT 870
 DB 201 MetAspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAsp 220
 QY 871 ATTTGGATATCTAAACATCTGTAACAGAGAGAAAGAGACTCATCTATGTGCACAATGAG 930
 DB 221 IleTrpIleSerAsnLeuValThrArgGluGluArgIleThrTyrValHisAsnGlu 240
 QY 931 CTAGCCAAACATCGAAGAAGATGCCAGATCAGCTGGAGTCGCTACTTTGTTCTCCAAAGAA 990
 DB 241 LeuAlaAsnMetGluGluAspProArgSerAlaGlyValAlaThrPheValLeuGlnGlu 260
 QY 991 GAATTTGATAGATATTCGGCTATTGGGTGCTCCAAAGCTGGAACCAACTCCCAAGGTT 1050
 DB 261 GluPheAspArgTyrSerGlyTyrTrpCysProGlnAlaGluArgThrProSerGly 280
 QY 1051 GGTAAATTTCTAGAAATCTATATCAAGAAATATGATGAATCTGAGGTGGAATTTATTCAT 1110
 DB 281 GlyLysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHis 300
 QY 1111 GTTACATCCCCTATGTTGGAAACAAGAGGGCAGATTTCCTTCCTTAAACACAGGT 1170
 DB 301 ValThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGly 320
 QY 1171 ACAGCAATCTCTAAAGTCACTTTTAAGATGTCAGAAATATATGATGATCTGAAGGAGG 1230
 DB 321 ThrAlaAsnProLysValThrPheLysMetSerGluIleValValAspAlaAlaGlyGly 340
 QY 1231 ATCATAGATGTCATAGATAAGGAATTAATTCACCTTTGAGATCTTATTTGAAGAGGATT 1290
 DB 341 IleIleAspValIleAspLysGluLeuValGlnProPheGluIleLeuPheGluGlyVal 360
 QY 1291 GAATATATTGCCAGAGCTGGATGCTCTCTGAGGGAATAATATGCTTGCTCATCTACTA 1350
 DB 361 GluTyrIleAlaArgAlaGlyTrpThrProGluGlyLysHisAlaTrpSerIleLeuLeu 380
 QY 1351 GATCGCTCCAGACTCGCTCAGATAGTGTGTGATCTCACCTGAATTTATTTATCCAGTA 1410
 DB 381 AspArgSerGlnThrHisLeuGlnIleValLeuIleSerProGluLeuPheIleProVal 400
 QY 1411 GAAGATGATGTTATGAAAGGAGGAGACTCATCTGATGCTGCTGCTGCTGCTGCTGCTG 1470
 DB 401 GluAspAlaMetAspArgGlnArgLeuIleGluSerValProAspSerValThrPro 420
 QY 1471 CTAATATTCTATCAAGAAACAACAGACATCTCGATAAATAATATCCATGACATCTTTCATGTT 1530
 DB 421 LeuIleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisVal 440
 QY 1531 TTTCCCCCAAGTCACGAAGGAAATTTAGTTTATTTTGGCTCTGAAATGCAAAACAGGT 1590
 DB 441 PheArgHisLeuTyrIleThrSerIleLeuLysGluSerLysTyrLysArgSerSer 460
 QY 1591 TTCGTCATTTATCAAAAATTACATCTATTTTAAAGGAAAGCAATATATAACGATCCAGT 1650
 DB 461 PheArgHisLeuTyrIleThrSerIleLeuLysGluSerLysTyrLysArgSerSer 480
 QY 1651 GGTGGCTGCTGCTCCCAAGTGTCAAGTGTCTTATCCAAAGAGAGATAGCAATATACC 1710
 DB 481 GlyGlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleThrIleThr 500
 QY 1711 AGTGGTCAATGGCAAGTTCTTGGCCGCGCATGATCTTAATATATCAAGTTGATGAAGTCAGA 1770
 DB 501 SerGlyGluTrpGluValLeuGlyArgHisGlySerAsnIleTrpValAspGluAlaArg 520
 QY 1771 AGGCTGGTATATTTTGAAGCCACCAAGACTCCCTTTTAAAGCATCCTGCTGATGATGTC 1830
 DB 521 LysLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisIleLeuTyrValThr 540
 QY 1831 AGTTACGTAAATCTCTGAGAGGTGACACAGCTGACCTGCGGTGCTACTACATCTTTGTC 1890
 DB 541 SerTyrAlaAsnProGlyGluValArgLeuThrAspArgGlyTyrSerHisSerCys 560
 QY 1891 TGCATCAGTCAGCACTGTGACTTCTTTTATAAGTAAAGTATAGTAAACCAAGAAATCCACAC 1950

Db	561	CysLeuSerA ^r GHisCysAspPheIleSerLysTy ^r SerAsnGlnLysAsnProHis	580
Qy	1951	TCTGTGTCCTTTACAAGCTATCAAGCTCTGAAGATGACCCAACTTGCAAACAAAGAA	2010
Db	581	CysValSerLeuTy ^r LysLeuSerSerProGluAspAspProValHisLysThrLysGlu	600
Qy	2011	TTTTGGGCCACCATTTTGGATTTCAGCAGGCTCTCTTCCTGACTATACCTCTCCAGAAATT	2070
Db	601	PheT ^r PaLaThrIleLeuAspSerAlaGlyProLeuProAspTy ^r ThrProGluIle	620
Qy	2071	TTCTCTTTGAAAGTACTACTGGATTTCACATTCCTATGGGATGCTCTACAAGCCCTCATGAT	2130
Db	621	PheSerPheGluSerThrGlyPheThrLeuTy ^r GlyMetLeuTy ^r LysProHisAsp	640
Qy	2131	CTACAGCCTGGAAAGAAATATCTTACCTGTGCTCTTTCATATATGGTGGTCTCAGGTGCAG	2190
Db	641	LeuGlnProGlyLysLysTy ^r ProThrValLeuPheIleTy ^r GlyProGlnValGln	660
Qy	2191	TTGGTGAATAATCGGTTTAAAGGAGTCAAGTATTTCGGCTTCAATACCTACGCTCTCTA	2250
Db	661	LeuValAsnAsnArgPheLysGlyValLysTy ^r PheArgLeuAsnThrLeuAlaSerLeu	680
Qy	2251	GGTTATGTGTTGTAGTATAGACACACAGGGGATCTGTCAACGAGGGCTTAAATTTGAA	2310
Db	681	GlyTy ^r ValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGlu	700
Qy	2311	GGCGCTTTAAATATAAAATGGGTCAAAATAGAATATGACGATCAGTGGAGGAGCTCAA	2370
Db	701	GlyAlaPheLysTy ^r LysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGln	720
Qy	2371	TATCTAGCTTCTCGATATGATTCATTCATTCAGTCTAGATCGTGGGATCCACGGCTGGTCC	2430
Db	721	TyrLeuAlaSerGlnTy ^r AspPheIleAspLeuAspArgValGlyIleHisGlyT ^r rpSer	740
Qy	2431	TATGGAGGATACCTCTCCCTGATGGCATTAATCAGAGGTCAGATATCTTCAGGTTGCT	2490
Db	741	TyrGlyGlyTy ^r LeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAla	760
Qy	2491	ATTGCTGGGCCCGAGTCACCTCTGTGATCTCTATGATCAGGATACACGGACGGTAT	2550
Db	761	IleAlaGlyAlaProValThrLeuIleIlePheTy ^r AspThrGlyTy ^r ThrGluArgTy ^r	780
Qy	2551	ATGGGTCACTCCGACGAGATGAACAGGGCTATTACTTAGGATCTGTGGCATTCGACCA	2610
Db	781	MetGlyHisProAspGlnAsnGlnGlyTy ^r LysGlySerValAlaMetGlnAla	800
Qy	2611	GAAGAATCCCTCTGAACCAAAATCGTTTACTGCTCTTACATGGTTTCTCGATGAGAT	2670
Db	801	GluLysPheProSerGluProAsnArgLeuLeuLeuLeuHisGlyPheLeuAspGluAsn	820
Qy	2671	GTCCATTTTGACATACACGATATATTACTGAGTCTTTTATGTGAGGCTGGAAAGCCATAT	2730
Db	821	ValHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTy ^r	840
Qy	2731	GATTTACAGATCTATCTCTCAGGAGAGACACAGCATAAGATGTTCTTGAATCCGGAGAACAT	2790
Db	841	AspLeuGlnIleTy ^r ProGlnGluArgHisSerIleArgValProGluSerGlyGluHis	860
Qy	2791	TATGAATGTCATCTTTTGCACTACCTTCAAGAAAACCTTGGATCAAGTATTCGCTCTA	2850
Db	861	TyrGluLeuHisLeuHisTy ^r LeuGlnGluAsnLeuGlySerArgIleAlaAlaLeu	880
Qy	2851	AAAGTGATA	2859
Db	881	LysValIle	883

DT	15-JUL-2003	(first entry)	
DE	Human protein modification and maintenance molecule-6 (PMM6-6).		
XX	Human; protein modification and maintenance molecule; PMM6; cancer;		
XX	cell proliferation disorder; atherosclerosis; neurological disorder;		
XX	epilepsy; Huntington's disease; stroke; immune disorder; allergy;		
XX	inflammatory disorder; AIDS; developmental disorder; hypothyroidism;		
XX	Cushing's syndrome; gastrointestinal disorder; epithelial disorder;		
XX	infection; cystostatic; antiarteriosclerotic; anticonvulsant; nootropic;		
XX	neuroprotective; cerebroprotective; anti-HIV; antiallergic; vulnary;		
XX	antiinflammatory; thyromimetic.		
OS	Homo sapiens.		
XX	WO2003031939-A2.		
PN	17-APR-2003.		
XX	11-OCT-2002; 2002WO-US032850.		
PD	12-OCT-2001; 2001US-0329689P.		
XX	25-OCT-2001; 2001US-0335703P.		
PR	08-NOV-2001; 2001US-0348887P.		
PR	29-NOV-2001; 2001US-0334145P.		
PR	06-DEC-2001; 2001US-0337451P.		
PR	14-DEC-2001; 2001US-0340584P.		
XX	(INCY-) INCYTE GENOMICS INC.		
FA	Ramkumar J, Gorvad AE, Baughn MR, Emerling BM, Yang J, Lee SY;		
XX	Tran UK, Becha SD, Duggan EM, Lee EA, Griffin JA, Li JX, Yue H;		
PI	Sprague JW, Hafalia AUA, Chawla NK, Lehr-Mason PM, Kabie AE, Yue H;		
PI	Marquis JP, Yao MG, Richardson TW, Tang TY, Jin P, Chien D;		
PI	Bhatia U, Burrill JD, Lee S, Blake JJ, Ho A, Zheng W;		
XX	NPI; 2003-430274/40.		
DR	WP-PSDB; ACA92421.		
XX	New human protein modification and maintenance molecules (PMM6), useful		
PT	for diagnosing, treating and preventing diseases or conditions associated		
PT	with the aberrant PMM6 expression e.g. cancer, atherosclerosis, or		
PT	infections.		
XX	Claim 1; Page 242-243; 31pp; English.		
PS	The present invention relates to the isolation of human protein		
XX	modification and maintenance molecules (PMM6), and the polynucleotide		
CC	sequences encoding them. A total of 40 PMM6 polypeptides (designated PMM6		
CC	-1 to PMM6-40) are disclosed. The sequences of the invention are useful		
CC	for diagnosing a condition or disease associated with the expression of		
CC	PMM6 in a subject, preparing a polyclonal or monoclonal antibody, and		
CC	generating an expression profile of a sample containing the		
CC	polynucleotides. The diseases or conditions associated with decreased		
CC	expression or overexpression of PMM6 are cell proliferation disorders		
CC	(e.g. cancer, atherosclerosis), neurological disorders (e.g. epilepsy,		
CC	Huntington's disease, stroke), immune/inflammatory disorders (e.g. AIDS,		
CC	allergies), developmental disorders (e.g. hypothyroidism, Cushing's		
CC	syndrome), gastrointestinal or epithelial disorders, and infections. The		
CC	PMM6 polypeptides or their fragments are useful in screening compounds		
CC	for effectiveness as agonists or antagonists of the polypeptides, or in		
CC	altering the expression of the target polynucleotide and compounds that		
CC	specifically bind to, or modulate the activity of the polypeptide.		
CC	ABU92021-ABU92060 represent the human PMM6 polypeptides of the invention		
XX	Sequence 831 AA;		
SQ	Alignment Scores:		
XX	Pred. No.:	0	Length: 831
XX	Score:	4397.50	Matches: 831
XX	Percent Similarity:	94.2%	Conservative: 0
XX	Best Local Similarity:	94.2%	Mismatches: 0
XX	Query Match:	79.2%	Indels: 51

DB:	6	Gaps:	1
US-10-825-632-2 (1-3120) x ABU92026 (1-831)			
QY	214	ATGGCAGCAGCAATGAAACAGACAGCTGGGTGGTGGATATTTGAAACTGGCGACTGT	273
Db	1	MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAaspCys	20
QY	274	GAGGAGAAATATTCAATCACAGGATGGCTTAATTTGGAGCCCTTTTATGTTGGACGGTAT	333
Db	21	GluGluasnIleGluSerGlnAaspArgProLysLeuGluProPheTyrValGluAargTyr	40
QY	334	TCCTGGAGTCAGCTTAAAAAGCTGTGCTGCCGATACAGAAAATATCATGTCGTACATGATG	393
Db	41	SerTrpSerGlnLeuLysLysLeuLeuAlaAaspThrArgLysTyrHisGlyTyrMetMet	60
QY	394	GCTAAGGCCACCATGATTTTCATGTTTGTGGAGGAATGATCCAGATGGACCTCATTTCA	453
Db	61	AlaLysAlaProHisAaspPheMetPheValLysArgAenAaspProAaspGlyProHisSer	80
QY	454	GACAGATCTATTACCTCTGCATGCTGCTGAGAACAGAGAAAATACACTGTTTATTCT	513
Db	81	AspArgIleTyrTyrLeuAlaMetSerGlyGluasnArgGluasnThrLeuPheTyrSer	100
QY	514	GAATTTCCCAAACTATCAATAGACAGCAGCTTTAATGCTCTCTTGGAGCCCTCTTTTG	573
Db	101	GluIleProLysThrIleAenArgAlaValLeuMetLeuSerTrpLysProLeuLeu	120
QY	574	GATCTTTTTCAGCAACACTGACATGGAATGATTTCTCGAGAGAGAACTATTAAAGA	633
Db	121	AspLeuPheGlnAlaThrLeuAaspTyrGlyMetTyrSerArgGluGluLeuLeuArg	140
QY	634	GAAGAAAAGCGAATCGAAGTCGGAATGCTCTTACGATATTCACCAAGGAAGTGA	693
Db	141	GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAaspTyrHisGlnGlySerGly	160
QY	694	ACATTTCTGTTTCAAGCCGCTAGTGAATTTATCACTGTAAGAAGTGGAGGCCCAAGGA	753
Db	161	ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAaspGlyGlyProGlnGly	180
QY	754	TTTACGCAACAACTTTAAGGCCAATCTAGTGGAAACTAGTTGTCCCAACATACGATG	813
Db	181	PheThrGlnGlnProLeuArgProAenLeuValGluThrSerCysProAenIleArgMet	200
QY	814	GATCCAAAATTTAGCCCGCTGATCCAGACTGGATTGCTTTTATACATAGCAACGATTT	873
Db	201	AaspProLysLeuCysProAlaAaspProAaspTrpIleAlaPheIleHisSerAenAaspIle	220
QY	874	TGGATATCTAACATCTGTAACCAAGAGAAAGGAGACTCACTTATGTGCACAATGAGCTA	933
Db	221	TrpIleSerAenIleValThrArgGluGluAargLeuThrTyrValHisAenGluLeu	240
QY	934	GCCAACTGGAAGAAGATGCCAGATCAGCTGGAGTCGCTTACTTTGTTCTCCAAGAGAA	993
Db	241	AlaAenMetGluGluAaspAlaAargSerAlaGlyValAlaThrPheValLeuGlnGluGlu	260
QY	994	TTTGATAGATATTCTGGCTATTGGTGGTCTCCAAAGCTGAAACAACTCCCGAGTGGGT	1053
Db	261	PheAaspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly	280
QY	1054	AAAACTCTAGAATTTCTATATGAAGAAAATGATGAATCTGAGGTGGAAATTTATTCATGTT	1113
Db	281	LysIleLeuArgIleLeuTyrGluGluAenAaspGluSerGluValGluIleIleVal	300
QY	1114	ACATCCCTATGTTGGAACAAGAGGGCAGATTCAATCCGTTATCTCTAAACAGGTACA	1173
Db	301	ThrSerProMetLeuGluThrArgArgAlaAaspSerPheArgTyrProLysThrGlyThr	320
QY	1174	GCAATCCTTAAGTCACTTTAAGATGTCAGAAAATGATGATGCTGAAGGAGGATC	1233
Db	321	AlaAenProLysValThrPheLysMetSerGluIleMetIleAaspAlaGluGlyArgIle	340
QY	1234	ATAGATGTCATAGTAAGGAACCTAAATTCACCTTTTGAGATTCTATTTGAAGGAGTTGAA	1293
Db	341	IleAaspValIleAaspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu	360
QY	1294	TATATTGCCAGCAGCTGGATGGACTCTCTGAGGGAAAATATGCTTGGTCCATCTACTAGAT	1353
Db	361	TyrIleAlaAargAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAasp	380
QY	1354	CGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATTTATTTATCCAGTAGAA	1413
Db	381	ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu	400
QY	1414	GATGATCTTATGAAAAGCAGAGACTCATTGAGTCAGTGCCTGATTTCTGTGACCCCACTA	1473
Db	401	AaspAaspValMetGluArgGlnArgLeuIleGluSerValProAaspSerValThrProLeu	420
QY	1474	ATTATCTATCAAGAAAACAACAGACATCTCGATAAATATCCATGACATCTTTTCATGTTT	1533
Db	421	IleIleTyrGluGluThrThrAepIleTyrIleAenIleHisAaspIlePheHisValPhe	440
QY	1534	CCCCAAAGTCCGAAAGGAAAATTGAGTTTATTTTTCCTCTGTAATGCAAAACAGGTTTC	1593
Db	441	ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe	460
QY	1594	CGTCAATTTATACAAATACATCTATTTTAAAGGAAGCAATATATAACGATCCAGTGGT	1653
Db	461	ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly	480
QY	1654	GGGTGCTCTCTCAAGTGAATTTCAAGTGTCTTATCAAGAGGAGATAGCAATACCAGT	1713
Db	481	GlyLeuProAlaProSerAaspPheLysCysProLysGluGluIleAlaIleThrSer	500
QY	1714	GGTGAATGGGAAGTTCTTGGCCGGCATGGATCTAATATCCAAAGTTGATGAAGTGAAGG	1773
Db	501	GlyGluTrpGluValLeuGlyArgHisGlySerAenIleGlnValAaspGluValArgArg	520
QY	1774	CTGTATATTTTGAAGCCCAAGACTCCCTTTAGAGCATCACCTGTAGTAGTCACT	1833
Db	521	LeuValTyrPheGluGlyThrLysAaspSerProLeuGluHisLeuTyrValValSer	540
QY	1834	TACTTAATCTCTGAGAGGTGACAAGCTGACTGACCGTGGTCTACTCACATTTCTTGCTGC	1893
Db	541	TyrValAenProGlyGluValThrArgLeuThrAaspArgGlyTyrSerHisSerCysCys	560
QY	1894	ATCAGTCAGCACTGTGACTTTCTTTTAAAGTATAGTAACTAGTAACTGCAAGAAATCCAC	1953
Db	561	IleSerGlnHisCysAaspPhePheIleSerLysTyrSerAenGlnLysAenProHisCys	580
QY	1954	GTGTCCCTTTTACAGCTATCAAGTCTCTGAAGATGACCCCACTTGCAAAACAAGAAATTT	2013
Db	581	ValSerLeuTyrLysLeuSerSerProGluAaspAaspProThrCysLysThrLysGluPhe	600
QY	2014	TGGGCCCACTTTTGGATTCCAGCAGTCTCTCTCTGACTATACCTCCAGAAATTTTC	2073
Db	601	TrpAlaThrIleLeuAaspSerAlaGlyProLeuProAaspTyrThrProGluIlePhe	620
QY	2074	TCTTTTGAAAGTACTACTGGAATTTACATTTGATGGATGCTCTACAAGCCTCATGATCTA	2133
Db	621	SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAaspLeu	640
QY	2134	CAGCTCGGAAGAAATATCTCTACTGTGCTGTTTCATATATATGTTGGTGGTCTCAGTGC	2193
Db	641	GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu	660
QY	2194	GTGTAATATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTCGCTCTCTAGGT	2253
Db	661	ValAenAenArgPheLysGlyValLysTyrPheArgLeuAenThrLeuAlaSerLeuGly	680
QY	2254	TATGTGTTGTAGTGTAGTACACACAGGATCTCTGTACCGAGGGCTTTAAATTTTGAAGC	2313
Db	681	TyrValValValValIleAenAenArgGlySerCysHisAargGlyLeuLysPheGluGly	700
QY	2314	GCCTTTAAATATAAATGGGTCAAAATAGAAATTTGACGATCAGGTGGAAGGACTCCAATAT	2373

Db 701 AlaPheLysTyLysMet----- 706
Qy 2374 CTAGCTTCTCGATGATGATTTTCATTTAGCTAGTCGTGGGCATCCACGGCTGCTCTAT 2433
Db 706 ----- 706
Qy 2434 GGAGGATACCTCTCCCTGATGGCATTATGCAGAGGTCAGATATCTTCAGGGTTCCTATT 2493
Db 707 -----ValAlaIle 709
Qy 2494 GCTGGGGCCCCAGTCACCTCTGTGATCTTCTATGATACAGGATACACGGACGTTATATG 2553
Db 710 AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet 729
Qy 2554 GGTCACTCCCTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAGACGAA 2613
Db 730 GlyHisProAspGlnAsnGluGlnGlyTyrLeuGlySerValAlaMetGlnAlaGlu 749
Qy 2614 AAGTTCCCTCTGAAACCAATCGTTTACTCTCTTACATGGTTTCTGTGATGAGATGTC 2673
Db 750 LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal 769
Qy 2674 CATTTTGCACATACCAGTATATTACTGAGTTTTTTAGTGAGGCTCGAAAGCCATATGAT 2733
Db 770 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 789
Qy 2734 TTACAGATCTATCTCTCAGGACAGACACAGATAGAGTTCTCTGAATCGGAGAACATTAT 2793
Db 790 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 809
Qy 2794 GAAGTCATCTTTTGCATCTTCAAGAAACCTTGGATCACTGATGCTGCTCTTAAAA 2853
Db 810 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeuLys 829
Qy 2854 GTGATA 2859
Db 830 ValIle 831

RESULT 10
ABU92030
ID ABU92030 standard; protein; 824 AA.
XX AC ABU92030;
XX DT 15-JUL-2003 (first entry)
XX DE Human protein modification and maintenance molecule-10 (PM9M-10).
XX KW Human; protein modification and maintenance molecule; PM9M; cancer;
KW cell proliferation disorder; atherosclerosis; neurological disorder;
KW epilepsy; Huntington's disease; stroke; immune disorder; allergy;
KW inflammatory disorder; AIDS; developmental disorder; hypothyroidism;
KW Cushing's syndrome; gastrointestinal disorder; epithelial disorder;
KW infection; cytostatic; antiatherosclerotic; anticonvulsant; nootropic;
KW neuroprotective; cerebroprotective; anti-HIV; anti-allergic; vulnerary;
KW anti-inflammatory; thymomimetic.
XX OS Homo sapiens.
XX PN WO2003031939-A2.
XX DX 17-APR-2003.
XX PF 11-OCT-2002; 2002WO-US032850.
XX PR 12-OCT-2001; 2001US-0329689P.
XX PR 25-OCT-2001; 2001US-0335703P.
XX PR 09-NOV-2001; 2001US-0348887P.
XX PR 28-NOV-2001; 2001US-0334345P.
XX PR 06-DEC-2001; 2001US-0337451P.
XX PR 14-DEC-2001; 2001US-0340584P.
XX PA (INCY-) INCYTE GENOMICS INC.

XX Ramkumar J, Gorvad AE, Baughn MR, Emerling BM, Yang J, Lee SY;
PI Tran UK, Becha SD, Duggan BM, Lee EA, Griffin JA, Li JX;
PI Sprague WW, Hafalia AJA, Chawla NK, Lehr-Nason PM, Kable AE, Yue H;
PI Marquis JP, Yao MG, Richardson TW, Tang YX, Jin P, Chien D;
PI Bhatia U, Burrill JD, Lee S, Blake JJ, Ho A, Zheng W;
XX WPI; 2003-430274/40.
DR N-PSDB; ACA92425.
XX New human protein modification and maintenance molecules (PM9M), useful
PT for diagnosing, treating and preventing diseases or conditions associated
PT with the aberrant PM9M expression e.g. cancer, atherosclerosis, or
PT infections.
XX Claim 1; Page 249-251; 31pp; English.
XX The present invention relates to the isolation of human protein
CC modification and maintenance molecules (PM9M), and the polynucleotide
CC sequences encoding them. A total of 40 PM9M polypeptides (designated PM9M
CC for PM9M-40) are disclosed. The sequences of the invention are useful
CC for diagnosing a condition or disease associated with the expression of
CC PM9M in a subject, preparing a polyclonal or monoclonal antibody, and
CC generating an expression profile of a sample containing the
CC polynucleotides. The diseases or conditions associated with decreased
CC expression or overexpression of PM9M are cell proliferation disorders
CC (e.g. cancer, atherosclerosis), neurological disorders (e.g. epilepsy,
CC Huntington's disease, stroke), immune/inflammatory disorders, (e.g. AIDS,
CC syndrome), developmental disorders (e.g. hypothyroidism, Cushing's
CC allergies), gastrointestinal or epithelial disorders, and infections. The
CC PM9M polypeptides or their fragments are useful in screening compounds
CC for effectiveness as agonists or antagonists of the polypeptides, or in
CC altering the expression of the target polynucleotide and compounds that
CC specifically bind to, or modulate the activity of the polypeptide.
CC ABU92021-ABU92060 represent the human PM9M polypeptides of the invention
XX SQ Sequence 824 AA;
Alignment Scores:
Pred. No.: 0 Length: 824
Score: 4353.00 Matches: 824
Percent Similarity: 93.4% Conservative: 0
Best Local Similarity: 93.4% Mismatches: 0
Query Match: 78.4% Indels: 58
DB: 6 Gaps: 1
US-10-825-632-2 (1-3120) x ABU92030 (1-824)
Qy 214 ATGCGAGCAGCAATGGAACAGACAGCTGGGTGTGAGATATTTGAACTCGGACTGT 273
Db 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
Qy 274 GAGGAGATATTGAATCAGAGATCGGCTAAATGGAGCCTTTTATGTTGACGGTAT 333
Db 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
Qy 334 TCCTGGAGTCAGCTTAAAGCTGCTGCCGATACAGAAATATCATGGCTACATGATG 393
Db 41 SerTrpSerGlnLeuLysLysLeuAlaPheThrArgLysTyrHisGlyTyrMetMet 60
Qy 394 GCTAAGGCACACATGATTTTCATGTTGTGAGAGGAATGATCCAGATGGACCTCATTC 453
Db 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
Qy 454 GACAGATCTATTACCTTGCCATGCTGGTGGAGACAGAGAAATACACTGTTTATTCT 513
Db 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
Qy 514 GAAATCCCAAACTATCAATAGAGCAGCTCTTAATGCTCTCTTGGAGCCTCTTTTG 573
Db 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerIleProLeuLeu 120
Qy 574 GATCTTTTTCAGGCACACATGGACTATGGAATGATTTCTCGAGAGAGAACTATTAGA 633

Db 121 ||||| AspLeuPheGln----- 124
Qy 634 GAAAGAAACGATGGAAACAGTCGGAATTGCTTTACGATTATCACCAAGGAAGTGA 693
Db 124 ----- 124
Qy 694 ACATTCTGTTTCAAGCCGGTAGTGAATTTATCAGCTAAAGATGGAGGCCACAAGGA 753
Db 124 ----- 124
Qy 754 TTACGCCAACACCTTTAAGGCCCAATCTAGTGAACACTAGTTGTCCCAACATACGGATG 813
Db 125 -----GlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 142
Qy 814 GATCCAAATATATGCCCGCTGATCCAGACTGGATTGCTTTTATACATAGCAACGATATT 873
Db 143 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 162
Qy 874 TGGATATCTAACATCGTTAACAGAGAAGAGAGACTCACTTATGTGCACAATGAGCTA 933
Db 163 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu 182
Qy 934 GCCAACATCGGAAGAAGATGCCAGATCGCTGAGTCGCTACCTTGTCTCCAAAGAA 993
Db 183 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu 202
Qy 994 TTGTATAGATATCTCGGCTATTGGTGGTGTCCAAAGCTGAACAACCTCCAGTGGTGT 1053
Db 203 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 222
Qy 1054 AAAATCTTAGAATCTATATGAGAAATGATCAATCTGAGGTGGAAATTTATTCATGTT 1113
Db 223 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 242
Qy 1114 ACATCCCTTATCTTGGAAACAAGAGGCGCAGATTCACTCCGTTATCTCAAAAACAGGTACA 1173
Db 243 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 262
Qy 1174 GCAATCTCAAAGTCACTTTAAGATGTGAGAAATATGATGATGCTGAGGAAGGATC 1233
Db 263 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 282
Qy 1234 ATAGATGTCATAGATAAGACTAAATCAACTTTGAGATCTCATTTGAAAGAGTTGAA 1293
Db 283 IleAspValIleAspLysGlnLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 302
Qy 1294 TATATTGCCAGAGCTGGATGGACTCTCGAGGGAATAATGCTTGGTCCATCTCTACTAGAT 1353
Db 303 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 322
Qy 1354 CGCTCCAGACTCGCTCAGATAGTGTGTGATCTCACTGAATATTATTTATCCAGTAGAA 1413
Db 323 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 342
Qy 1414 GATGATGTTATGGAAGGAGAGACTCAATTGAGTCAGTCCCTGATTTCTGTGACGCCACTA 1473
Db 343 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 362
Qy 1474 ATTATCTATGAGAAACACACACATCTGGATAATATCCATGACATCTTTTCATGTTTTT 1533
Db 363 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 382
Qy 1534 CCCAAAGTCACGAAGAGAAATTGAGTTTATTTTGGCTCTGAAATCAAAACAGGTTTC 1593
Db 383 ProGlnSerHisGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 402
Qy 1594 CGTCATTTATACAAATTCATCTATTTTAAAGGAACCAATATTAACGATCCAGTGGT 1653
Db 403 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 422
Qy 1654 GGCTCGCTGCTCCAGTCAATTTCAAGTGTCTTCAAGAGGAGATAGCAATTTACCAGT 1713
|||||

Db 423 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 442
Qy 1714 GGTGAATGGGAAGTTCTTTGGCGGCATGGATCTAATATTCGAAGTTGATGAAGTCAGAAG 1773
Db 443 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 462
Qy 1774 CTGGTATATTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTACGTAGTCAGT 1833
Db 463 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisIleLysLeuTyrValValSer 482
Qy 1834 TAGCTAAATCTCGAGAGGTGACAAGGCTGACACCGTCTGCTACTCACATTTCTTGTGTC 1893
Db 483 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 502
Qy 1894 ATCAGTCAGCACTGTGACTTCTTTATAAGTAAGTATAGTAAACAGAGAATCCACACTGT 1953
Db 503 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 522
Qy 1954 GTGTCCCTTTACAGCTATCAAGTCTCGAAGTGAACCACTTCGCAAAACAAGCAATTT 2013
Db 523 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 542
Qy 2014 TGGGCCACCATTTTGGATTACAGCTCTCTCTCTGCTGACTATATCTCTCCAGAAATTTTC 2073
Db 543 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePhe 562
Qy 2074 TCTTTTGAAGTACTACTCGATTATACATTTGTATGGATGCTCTACAAGCCTCATGATCTA 2133
Db 563 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 582
Qy 2134 CAGCTTGGAAAGAAATATCTCTGCTGTGTTTCATATATATGTTGGTCTCAGGTGAGTTG 2193
Db 583 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 602
Qy 2194 GTGAATATTCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTACCTGCTCTAGGT 2253
Db 603 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 622
Qy 2254 TATGTGTTGTAGTATAGACACACAGGGGATCTCTCTACCGAGGCTTTAAATTTTGAAGGC 2313
Db 623 TyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 642
Qy 2314 GCCTTTAAATATAAATGGGTCAATAGAAATTTGACGATCAGTGGGAAGGACTCCAATAT 2373
Db 643 AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr 662
Qy 2374 CTAGCTTCTCGATATGATTTTCATTTAGATCTGTGGGATCCCGGTCCTGCTCTAT 2433
Db 663 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyr 682
Qy 2434 GGAGGATACCTCTCCCTGATGGCATTAAATGACAGAGTCAGATATCTTCAGGGTTGCTATT 2493
Db 683 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 702
Qy 2494 GCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAAGCTTATATG 2553
Db 703 AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet 722
Qy 2554 GGTCAACCTCGACAGAAATGAACAGGCTTATTTAGGATCTGTGGCCATCCAGACAGAA 2613
Db 723 GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 742
Qy 2614 AAGTTCCCTCTGAACCAATCGTTTACTGCTTCTACATGTTTCTCTGATGAGAAATGTC 2673
Db 743 LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal 762
Qy 2674 CATTTTGCACATACCATATTTACTGAGTTTTTTAGTGAGGCTCGGAAGCCATATCAT 2733
Db 763 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 782
Qy 2734 TTAAGATCTATCTCTCAGGAGACACAGCAATAGAGTTCCTGAAATCGGGAACAATTAT 2793
Db 783 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 802
|||||

QY 2794 GAACATGCACTTTTGCACCTACCTTCAAGAAACCTTGGATCAGGTATTCCTCTAAAAA 2853
 DB 803 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnGlySerArgIleAlaLeuLys 822

QY 2854 GTGATA 2859
 DB 823 ValIle 824

RESULT 11
 ID ABB97361 standard; protein; 782 AA.
 XX ABB97361;
 XX ABB97361;
 XX 27-JUN-2002 (first entry)
 XX Novel human protein SEQ ID NO: 629.
 DE Human; anti-anemic; vulnerary; anti-inflammatory; immunomodulator;
 KW antifertility; cerebroprotective; cytosstatic; rheumatic; gene therapy;
 KW neuroprotective; antiparkinsonian; protein therapy; EST;
 KW expressed sequence tag.
 XX Homo sapiens.
 OS WO200222660-A2.
 PN 21-MAR-2002.
 XX 10-SEP-2001; 2001WO-US026015.
 XX 11-SEP-2000; 2000US-00659671.
 PR (HYSE-) HYSEQ INC.
 PA Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao Qa, Ren F;
 PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;
 DR WPI: 2002-292408/33.
 DR N-PSDB; ABN32547.
 XX An isolated polynucleotide for treating diseases associated with its
 PT encoded polypeptide such as cancer and multiple sclerosis.
 PS Example 2; SEQ ID NO 629; 509pp; English.
 XX The present invention provides the protein and coding sequences of 444
 CC human proteins. These were isolated from expressed sequences tags
 CC (ESTs). They can be used to stimulate cell growth, to regulate regrowth
 CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
 CC e.g. in burn treatment, to regulate the immune system e.g. to treat
 CC multiple sclerosis, to regulate activin or inhibin e.g. to treat
 CC infertility, to regulate haemostasis or thrombolysis e.g. to treat stroke
 CC and cancer, to screen for drugs, to treat inflammatory conditions e.g.
 CC rheumatoid arthritis, and to treat nervous system disorders e.g.
 CC Parkinson's disease. The present sequence is a protein of the invention
 XX
 SQ Sequence 782 AA;

Alignment Scores:
 Pred. No.: 0 Length: 782
 Score: 4118.00 Matches: 782
 Percent Similarity: 88.7% Conservative: 0
 Best Local Similarity: 88.7% Mismatches: 0
 Query Match: 74.2% Indels: 100
 DB: 5 Gaps: 1

US-10-825-632-2 (1-3120) x ABB97361 (1-782)

QY 214 ATGGCAGCAGCAATCGAAACAGACAGCTGGGTCTTGAGATATTTGAACCTGGGACTCT 273
 DB 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20

QY 274 GAGGAGATATTGAATCAAGATCGGCCTAAATTTGAGCCCTTTTATGTTGACCGGTAT 333
 DB 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40

QY 334 TCTGGAGTCAGCTTAAAGAGCTGCTCCGATACCGAGAAATATCATGCTCATCATGATG 393
 DB 41 SerTrpSerGlnLeuLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60

QY 394 GCTAAGGCACCATGATGTTTGTGTGAAGGAATGATCCAGATGAGCACTCATTCATCA 453
 DB 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80

QY 454 GACAGAAATCTATTACCTTGCCTATGCTGTCGAGAACAGAGAAAAATACACTGTTTTATTCT 513
 DB 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100

QY 514 GAAATTCCTCAAAATCTCAATAGACGACGCTCTTAATGCTCTCTTGGAGCCCTCTTTTG 573
 DB 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120

QY 574 GATCTTTTTCAGCAACACTGGACTATGGAATGTATTCTCGAGAGAGAACTATTAAAGA 633
 DB 121 AspleuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140

QY 634 GAAAGAAACGCAATTCGACAGCTCGGAATTCCTTACGATATCACCAAGAGAGTGA 693
 DB 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160

QY 694 ACATTTCTGTTTCAAGCCGCTAGTGAATTTATCAGCTAAAGATGGAGGCCACACAGA 753
 DB 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180

QY 754 TTTACGCAACAACTTTAAGCCCAATCTAGTGGAACTAGTCTCCCAACATACGAGATG 813
 DB 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200

QY 814 GATCCAAAATATTATGCCCGCTGATCCAGACTGGATTCGCTTTATATACATACAGCATATT 873
 DB 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220

QY 874 TGATATCTAAACATCGTAACACAGAGAGAAAGAGACTCACTTATGTGCACATGAGCTA 933
 DB 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu 240

QY 934 GCCAACATGGAAGAAGATGCCAGATCAGCTGGAGTCGCTACTTTGTTCTCCAGAGAGAA 993
 DB 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu 260

QY 994 TTTGATAGATATTCTGGCTATTGGGTGTCCTCAAAAGCTGAAACAACTCCCGAGTGGTGGT 1053
 DB 261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280

QY 1054 AAAATTTTAGATTTCTATGACAAAATGATGAATCTGAGGTGGAAATTTATTCATGTT 1113
 DB 281 LysIleLeuArgIleLeuTyrGluIleAsnAspGluSerGluValGluIleHisVal 300

QY 1114 ACATCCCTTATGTTGGAAACAGAGGGCAGATTTCCTTCCTTAAACAGAGTACA 1173
 DB 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320

QY 1174 GCAAAATCTTAAAGTCATTTTAAAGATGTCAGAAATATGATGTGCTGAAGAGAGGATC 1233
 DB 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340

QY 1234 ATAGATGTCATAGATAAGGAACATAATTCACCTTTTGAGATTCTATTGAAGAGATTGAA 1293
 DB 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360

QY 1294 TATATTGCCAGAGCTGGATGGACTCTCTGAGGGAATAATGCTTGGTCCATCTCTAGAT 1353
 DB 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380

1354 QY CGCTCCAGACTCGCTACAGATAGTGTTCATCTCACCTGAATATTATATCCAGTAGAA 1413
1355 DB |||||-----ValAlaIle 660
381 DB ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
1414 QY GATGATGTTATGGAAGCAGAGACTCATTTGAGTCAGTGCCTGATTTCTGTGACGCCACTA 1473
1415 DB |||||-----ValAlaIle 660
401 DB AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
1474 QY ATTATCTATGGAAGAAACAACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTTT 1533
1475 DB |||||-----ValAlaIle 660
421 DB IleIleIleThrGluThrThrAspIleIleIleAsnIleHisAspIlePheHisValPhe 440
1534 QY CCCAAATCTACGAAGAGAAATTTAGTTTATTTTTCCTCTCAATGCAAAACAGATTTC 1593
1535 DB |||||-----ValAlaIle 660
441 DB ProGlnSerHisGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
1594 QY COTCATTTATCAAAATTTACATCTATTTTAAAGAAAGCAAAATATAAGATCCAGTGGT 1653
1595 DB |||||-----ValAlaIle 660
461 DB ArgHisLeuThrLysIleThrSerIleLeuLysGluSerLysThrLysArgSerGly 480
1654 QY GGGCTGCTGCTCCCAAGTGATTTCAAGTGCTCTATCAAGAGAGATAGCAATTTACCACT 1713
1655 DB |||||-----ValAlaIle 660
481 DB GlyLeuProAlaProSerAspPheLysCysProIleLysGluIleAlaIleThrSer 500
1714 QY GGTGAATGGGAAGTTCTTTGGCCGCATGGATCTAATATCCAACTTTCATGAAGTCAGAGG 1773
1715 DB |||||-----ValAlaIle 660
501 DB GlyGluThrGluValLeuGluArgHisGlySerAsnIleGlnValAspGluValArgArg 520
1774 QY CTGGTATATTTGAAGGCAACAAAGACTCCCTTTAGAGCATCACCTGTACCTGATGTCAGT 1833
1775 DB |||||-----ValAlaIle 660
521 DB LeuValThrPheGluGluThrLysAspSerProLeuGluHisLeuThrValValSer 540
1834 QY TACGTAATCTCGGAGGTCACAGCTGACTGACCGTGCTACTCACATTTCTGCTGC 1893
1835 DB |||||-----ValAlaIle 660
541 DB TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyThrSerHisCysCys 560
1894 QY ATCAGTCAGACTGTGACTTCTTTAAGTAAGTATAGTAAACCAAGAGATCCACACTGT 1953
1895 DB |||||-----ValAlaIle 660
561 DB IleSerGlnHisCysAspPheIleSerLysThrSerAsnGlnLysAsnProHisCys 580
1954 QY GTGCTCCCTTTACAAGCTATCAAGTCTGAAGATCAACCACTGCTGCAAAACAAAGGAATTT 2013
1955 DB |||||-----ValAlaIle 660
581 DB ValSerLeuThrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
2014 QY TGGGCCACATTTTGGATTCAGCAGGTCTCTCTGACTATATCTCTCCAGAAATTTTC 2073
2015 DB |||||-----ValAlaIle 660
601 DB TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspThrProProGluIlePhe 620
2074 QY TCTTTTGAAGTACTACTGGATTATCATTTGATGGATGCTCTACAGCCTCATGATCTA 2133
2075 DB |||||-----ValAlaIle 660
621 DB SerPheGluSerThrThrGlyPheThrLeuThrGlyMetLeuThrLysProHisAspLeu 640
2134 QY CAGCTTGGAAAGAAATATCTATCTGTCTGTTCATATATGTTGCTCTCAGGTGCAGTTG 2193
2135 DB |||||-----ValAlaIle 660
641 DB GlnProGlyLysLysThrProThrValLeuPheIleThrGlyGlyProGln----- 657
2194 QY GTGAATAATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCTTAGCTCTTAGGT 2253
2195 DB |||||-----ValAlaIle 660
657 DB ----- 657
2254 QY TATGTGTTGTAGTAGACACACAGGGGATCTGTCCACCGAGGGCTTAAATTTGAAGGC 2313
2255 DB |||||-----ValAlaIle 660
657 DB ----- 657
2314 QY GCCTTTAAATATAAAATGGGTCAAATAGAAATTCACGATCAGTGGAGGACTCCAATAT 2373
2315 DB |||||-----ValAlaIle 660
657 DB ----- 657
2374 QY CTAGCTTCTCGATGATGATTTTCATTTGACTTTAGATCTGTGGGCATCCACGGCTGCTCTAT 2433
2375 DB |||||-----ValAlaIle 660
657 DB ----- 657
2434 QY GGAGGATACCTCTCCCTGATGGCATTAATGACAGAGTTCAGATATCTTCAGGGTTGCTATT 2493

658 DB -----ValAlaIle 660
2494 QY GTTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACGTATATG 2553
2495 DB |||||-----ValAlaIle 660
661 DB AlaGlyAlaProValThrLeuThrPheThrAspThrGlyThrThrGluArgTy-Met 680
2554 QY GGTCACTGACACAGATGAACAGGGCTATTACTAGATCTGTGGCCATGCAAGCAGAA 2613
2555 DB |||||-----ValAlaIle 660
681 DB GlyHisProAspGlnAsnGluGlnGlyThrThrLeuGlySerValAlaMetGlnAlaGlu 700
2614 QY AGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGTTTCTCTGGATGAGAATGTC 2673
2615 DB |||||-----ValAlaIle 660
701 DB LysPheProSerGlnProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal 720
2674 QY CATTTTGCATACACAGATATATTACTGATGTTTTTTAGTGGGTGGAAGCCATATGAT 2733
2675 DB |||||-----ValAlaIle 660
721 DB HisPheAlaHisThrSerIleLeuLeuSerPheLeuValAlaGlaGlyLysProThrAsp 740
2734 QY TTTACAGATCTATCCTCAGGAGACACAGCATATAGAGTTCTGATCGGAGAACATAT 2793
2735 DB |||||-----ValAlaIle 660
741 DB LeuGlnIleThrProGlnGluHisThrLeuGlnAsnLeuGlySerGlyGluHisThr 760
2794 QY GAATCGATCTTTTGCACCTACCTTCAAGAAACCTTGGATCACGTTATGCTCTTAA 2853
2795 DB |||||-----ValAlaIle 660
761 DB GluLeuHisLeuLeuHisThrLeuGlnAsnLeuGlySerArgIleAlaLeuLeuLys 780
2854 QY GTGATA 2859
781 DB |||||-----ValAlaIle 660
782 DB ValIle 782
RESULT 12
AD116334
ID AD116334 standard; protein; 746 AA.
XX AD116334;
AC AD116334;
XX
DT 22-APR-2004 (first entry)
XX
DE Human protein modification and maintenance molecule (PMM) protein #19.
XX
KW human; protein modification and maintenance molecule; PMM;
KW gastrointestinal disorders; peptic ulcer; Crohn's disease;
KW cardiovascular disorders; hypertension; congenital heart disease;
KW autoimmune disease; inflammatory disease; AIDS; anaemia;
KW developmental disorder; Cushing's syndrome; tubular acidosis;
KW epithelial disorder; eczema; scabies; neurological disorder;
KW Alzheimer's disease; multiple sclerosis; infection; cancer.
XX
OS Homo sapiens.
XX
XX WO2003100016-A2.
XX
PD 04-DEC-2003.
XX
PF 22-MAY-2003; 2003WO-US016498.
XX
PR 22-MAY-2002; 2002US-0383491P.
PR 24-JUN-2002; 2002US-0391378P.
PR 22-JUL-2002; 2002US-0397921P.
XX
PA (INCY-) INCYTE CORP.
XX
PI Khare R, Bulloch SA, Swarnakar A, Elliott VS, Marquis JP;
PI Mason PM, Chawla NK, Rankumar J, Kable AE, Hafalia AJA, Lee SY;
PI Tran UK, Yue H, Becha SD, Griffin JA, Chang H, Jiang X, Jackson AA;
PI Richardson TW, Lal PG, Yao MG, Lu Y, Warren BA, Jin P, Wilson AD;
PI Gietzen KJ;
XX
XX WPI; 2004-035124/03.
DR N-PSDB; AD116386.
XX
PT New protein modification and maintenance molecules, useful for diagnosing

PT or treating e.g. peptic ulcer, hypertension, rheumatic fever, AIDS, or
PT Cushing's syndrome, Alzheimer's disease, multiple sclerosis, stroke or
PT cancers.

XX Claim 1: SEQ ID NO 19; 419pp; English.

XX The invention comprises the amino acid and coding sequences of human
CC protein modification and maintenance molecules (PMM). The DNA and
CC protein sequences of the invention are useful for the diagnosis and
CC treatment of disorders associated with expression of PMM, such as:
CC gastrointestinal disorders (e.g. peptic ulcer and Crohn's disease),
CC cardiovascular disorders (e.g. hypertension and congenital heart
CC disease), autoimmune or inflammatory disease (e.g. AIDS and anaemia),
CC developmental disorders (e.g. Cushing's syndrome and tubular acidosis),
CC epithelial disorders (e.g. eczema and scabies), neurological disorders
CC (e.g. Alzheimer's disease and multiple sclerosis), infections and cancer.
CC The present amino acid sequence represents a human PMM protein of the
CC invention.

XX SQ Sequence 746 AA;

Alignment Scores:

Pred. No.:	0	Length:	746
Score:	2901.00	Matches:	746
Percent Similarity:	84.6%	Conservative:	0
Best Local Similarity:	84.6%	Mismatches:	0
Query Match:	70.3%	Indels:	136
DB:	6	Gaps:	1

US-10-825-632-2 (1-3120) x ADI16334 (1-746)

QY	214	ATGGCAGCAGCATGGAAACAGAACAGCTGGGTGGTGGATATTTGAACTCGGAGCTGT	273
DB	1	MetAlaAlaMetGluThrGluGluLeuGlyValGluLeuPheGluThrAlaAspCys	20
QY	274	GAGGAAATATTGAATCAGAGATCGCCTAAATTTGGAGCTTTTATGTGAGCGGTAT	333
DB	21	GluGluAenilleGluSerGlnAspArgProLysLeuGluProPheYrValGluArgYr	40
QY	334	TCCTGGAGTCAGCTTAAAGCTGCTTGGCGATACACGAAATATCATCGCTACATGATG	393
DB	41	SerTrpSerGlnLeuLysLysLeuAlaAspThrArgLysYrHisGlyYrMetMet	60
QY	394	GCTAAGGCACCATGATTTTCATGTTGTGAAGAGGAATGATCCAGATGACCTCATTTCA	453
DB	61	AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer	80
QY	454	GACAGATCTATTACCTTGCCATGCTGCTGGTGAGAACAGAGAAATACACTGTTTATTCT	513
DB	81	AspArgLleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheYrSer	100
QY	514	GAATTTCCNAACTCATATAGACGACGCTTATATGCTCTTGGAGGCTCTTTTG	573
DB	101	GluileProLysThrileAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu	120
QY	574	GATCTTTTTCAGGCAACTCGACTATGGAATGATTTCTCGAGAAGACATATTAGA	633
DB	121	AspLeuPheGlnAlaThrLeuAspYrGlyMetYrSerArgGluGluLeuLeuArg	140
QY	634	GAAGAAACCGATTTGGAACAGTCGGAATTCCTTTACGATTATCACCAAGAAAGTGA	693
DB	141	GluArgLysArgileGlyThrValGlyleAlaSerYrAspYrHisGlnGlySerGly	160
QY	694	ACATTTCTGTTTCAAGCCGGTAGTGGAAATTTATCACGTAAAGATGGAGGCGCACAGGA	753
DB	161	ThrPheLeuPheGlnAlaGlySerGlyileYrHisValLysAspGlyGlyProGlnGly	180
QY	754	TTTACGCAACACCTTTTAAGGCCCAATCTAGTGGAAACTAGTTGTCCCAACATACGATG	813
DB	181	PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerYrSerProAenilleArgMet	200
QY	814	GATCCAAATATATGCCCGCTGATCCAGACTGGAATTCCTTTATACATAGCAACGATATT	873

DB	201	AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle	220
QY	874	TGGATATCTTAACATCGTAACAGAGAGAGAGAGACTCCTATTATGTGCAATAGCTA	933
DB	221	TrpIleSerAsnilleValThrArgGluGluArgGluLeuThrYrValHisAsnGluLeu	240
QY	934	GCCAACTGGAAGAAGATGCCAGATCAGCTGAGTCCGTACTCTTTGTTCTCCAAAGAA	993
DB	241	AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu	260
QY	994	TTTGATAGATATTCGTGGCTATTGGTGTCCAAAGCTGAAACAACTCCAGTGGTGGT	1053
DB	261	PheAspArgYrSerGlyYrTrpTrpCysProLysAlaGluThrThrProSerGlyGly	280
QY	1054	AAAAATCTTAGAATCTATATGAGAAATGATGAATCTGAGGTGGAAATATTATCATGTT	1113
DB	281	LysileLeuArgileLeuYrGluGluAsnAspGluSerGluValGluIleHisVal	300
QY	1114	ACATCCCTTATGTTGMAACAGAGGAGGAGATTCATTCCTGTTATCTAAACACAGTACA	1173
DB	301	ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgYrProLysThrGlyThr	320
QY	1174	GCAATCTCTAAAGTCACTTTTAAAGATGTCAAAATATATGATTGATCTGGAAGAGATC	1233
DB	321	AlaAsnProLysValThrPheLysMetSerGluileMetileAspAlaGluGlyArgIle	340
QY	1234	ATAGATGTCATAGTAAGAACTAATTCACCTTTTTCAGATTTCTATTGGAAGGATGAA	1293
DB	341	IleAspValilleAspLysGluLeuileGlnProPheGluileLeuPheGluGlyValGlu	360
QY	1294	TATATTGCCAGAGCTGGATGGACTCTCTGAGGAAATATGCTGGTCCATCTCTACTAGAT	1353
DB	361	TyrileAlaArgAlaGlyYrThrProGluGlyYrLysYrLysYrLysYrLysYrLysYr	380
QY	1354	CGCTCCAGACTCGGCTCACAGATGTTGATCTCACTGAAATATTATTATCCCAAGTAA	1413
DB	381	ArgSerGlnThrArgLeuGlnilleValleuileSerProGluLeuPheileProValGlu	400
QY	1414	GATGATGTTATGGAAGGCAGAGACTCATTCAGTCAGTGCCTGATCTGTGACCCACTA	1473
DB	401	AspAspValMetGluArgGlnArgLeuileGluSerValProAspSerValThrProLeu	420
QY	1474	ATTATCTATGAAGAAACACAGACATCTGGATAATATCCATGACATCTTTTCATGTTT	1533
DB	421	IleileYrGluGluThrThrAspIleTrpIleAsnilleHisAspIlePheHisValPhe	440
QY	1534	CCCCAAAGTCACGAGAGGAAATTCAGTTTATTTTTCCTCTGATCTGAAATGCAAAACAG	1593
DB	441	ProGlnSerHisGluGluGluileGluPheilePheAlaSerGluCysLysThrGlyPhe	460
QY	1594	CGTCATTTATACAAATACATCTATTTTAAAGGAAAGCAATATATAACCATCCAGTGGT	1653
DB	461	ArgHisLeuYrLysileThrSerileLeuLysGluSerLysYrLysYrLysYrSerGly	480
QY	1654	GGCTGCTGCTCCAGTGATTTCAAGTCTCTATCAAGAGGAGATAGCAATATACCAGT	1713
DB	481	GlyLeuProLysProSerAspPheLysCysProLysGluGluileAlaileThrSer	500
QY	1714	GGTGAATGGGAAGTCTTCGCGCGCATCGATCTAATATCCCAAGTGTGATGAGTCAGAG	1773
DB	501	GlyGluTrpGluValLeuGlyArgHisGlySerAsnilleGlnValAspGluValArgArg	520
QY	1774	CTGATATATTTGAAAGGCACCAAGACTCCCTTTTAGAGCATCACCTGTAGCTAGTCACT	1833
DB	521	LeuValYrPheGluGlyThrLysAspSerProLeuGluHisHisLeuYrValIser	540
QY	1834	TACGTAATCTCGAGAGGTGCAAGGCTGACTGACCGTGGCTACTCACATTTCTGTCTGC	1893
DB	541	TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyYrSerHisSerCysCys	560
QY	1894	ATCAGTCAGACTGTGACTTCTTTTATAAGTAAGTATAGTACCAAGAGAAATCCCACTGT	1953
DB	561	IleSerGlnHisCysAspPhePheileSerLysYrSerAsnGlnLysAsnProHisCys	580


```
QY 1954 GTGTCCTTTCAAGCTATCAAGTCTCAAGATGACCCAACTTGCAAAACAAAGGAATTT 2013
DB 581 ValSerLeuTyrLysLeuSerProGluAspProThrCysLysThrLysGluPhe 600
QY 2014 TGGGCGACATTTGGATTACAGAGTCTCTTCCTGACTATATCTCTCAGAAATTTTC 2073
DB 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 620
QY 2074 TCTTTTGAAGTACTACTGGATTACATTTATGGGATGCTCTACAGGCTCATGATCTA 2133
DB 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
QY 2134 CAGCTCGGAAGAAATATCTACTGCTGTGTTTCATATATGTTGCTCTCAGGTGCGAGTTG 2193
DB 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 660
QY 2194 GTGAATAATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCCCTAGCCTCTCAGGT 2253
DB 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680
QY 2254 TATGTGTTGTAGTATAGACACACAGGGGATCCTGTCCAGGGCTTAAATTTGAAGC 2313
DB 681 TyrValValValValIleAspAsnArgLysSerCysHisArgGlyLeuLysPheGluGly 700
QY 2314 GCCTTTTAAATATAAATCGGTCAAATAGAAATTCAGCATCAGGTGGAAGGACTCCCAATAT 2373
DB 701 AlaPheLysTyrLysMet----- 706
QY 2374 CTAGCTTCTCGATATGATTTCACTTGACTTAGATCGTGTGGGCATCCACGGCTGCTAT 2433
DB 706 ----- 706
QY 2434 GGAGGATACCTCTCCCTGATGGCATTATGCAGAGGTCAGATATCTTCAGGGTTGCTATT 2493
DB 706 ----- 706
QY 2494 GCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGATACACGGAACTTATATG 2553
DB 706 ----- 706
QY 2554 GGTCAACCTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAA 2613
DB 706 ----- 706
QY 2614 AAGTTCCCTCTGAACCAAAATCGTTTACTGCTCTTACATGTTTCTCGGATGAGATGTC 2673
DB 706 ----- 706
QY 2674 CATTTTGCATACACGATATATTACTGAGTTTTTTAGTGGGCTGGAAGCCATATGAT 2733
DB 706 ----- 706
QY 2734 TTACAGATCTATCTCAGGAGACACAGCATAAGAGTTCTGTAATCGGAGAACATTAT 2793
DB 707 -----IleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 724
QY 2794 GAAGTGCATCTTTTGCACTACCTTCAAGAAAACCTTGGATCACGATATTGCTGCTCAAAA 2853
DB 725 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLys 744
QY 2854 GTGATA 2859
DB 745 ValIle 746
RESULT 13
ADT04044
ID ADT04044 standard; protein; 738 AA.
XX
AC ADT04044;
XX
DT 30-DEC-2004 (first entry)
XX
```

```
DE Human protein modification and maintenance molecule protein SeqID24.
XX protein modification and maintenance molecule; PMMM; cytostatic;
KW immunomodulator; expression; immune disorder; cancer; human.
XX Homo sapiens.
XX WO2004084806-A2.
XX 07-OCT-2004.
XX 16-MAR-2004; 2004WO-US0080006.
XX 21-MAR-2003; 2003US-0456864P.
PR 03-APR-2003; 2003US-0460512P.
PR 19-MAY-2003; 2003US-0472027P.
PR 30-MAY-2003; 2003US-0475072P.
PR 30-MAY-2003; 2003US-0475190P.
PR 24-JUN-2003; 2003US-0482575P.
PR 13-AUG-2003; 2003US-0495151P.
XX (INCY-) INCYTE CORP.
XX Baughn MR, Marquis JP, Kable AE, Chawla NK, Emerling BM, Lee SY;
PI Hafalia AJA, Ramkumar J, Richardson TW, Wang JT, Nakamura LL;
PI Yang J, Jin P, Becha SD, Wilson AD;
DR WPI; 2004-710256/69.
DR N-PSDB; ADT04072.
XX New human protein modification and maintenance molecules (PMMM)
PT polypeptide, useful in preparing a composition for treating a disease
PT associated with decreased expression or overexpression of PMMM e.g.,
XX cancer.
XX Claim 1; SEQ ID NO 24; 222pp; English.
XX This invention relates to novel human protein modification and
CC maintenance molecule (PMMM) proteins and the DNA sequences which encode
CC them. The invention may be useful for the production of compounds with a
CC cytostatic or immunomodulator activity. The proteins are useful in
CC preparing a composition for diagnosing or treating a disease or condition
CC associated with decreased expression or overexpression of functional
CC PMMM, for example immune disorders or cancer. The present sequence is
CC that of a human protein modification and maintenance molecule (PMMM)
CC protein of the invention.
XX
SQ Sequence 738 AA;
Alignment Scores:
Pred. No.: 0 Length: 738
Score: 3817.00 Matches: 732
Percent Similarity: 83.1% Conservatives: 1
Best Local Similarity: 83.0% Mismatches: 5
Query Match: 68.8% Indels: 144
DB: 8 Gaps: 3
US-10-825-632-2 (1-3120) x ADT04044 (1-738)
QY 214 ATGGCAGCAGCAATGGAACACAGACAGCTGGGTGTGAGATATTTGAACTGCGGACTGT 273
DB 1 MetAlaAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
QY 274 GAGGAGAATATTGAATCACAGGATCGGCTTAATTTGGAGCCCTTTTATGTTGAGCGGTAT 333
DB 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
QY 334 TCCTGGAGTCAGCTTAAAGAGCTGTCGCCGATACAGAAAATATCATGGCTACATGATG 393
DB 41 SerTrpSerGlnLeuLysLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
QY 394 GCTAAGGCACCATCATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGGACTCATTC 453
```

Db 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
QY 454 GACAGAACTATTACCTTGCATGCTGGTGAGAACAGAGAAAATACACTGTTTATTCT 513
Db 81 AspArgIleTyrTyrLeuAlaMetSerGlyLysAsnArgGluAsnThrLeuPheTyrSer 100
QY 514 GAAATTCACAAACTATCAATAGAGACGACAGTCTTAATGCTCTCTTGGAGGCTCTTTTG 573
Db 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
QY 574 GATCTTTTTCAGCAACACTGGACTATGGAATGATTCTCGAAGAAGAACTATTAAAGA 633
Db 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140
QY 634 GAAAGAAACGATTTGGAACAGTCGGAATGCTCTTACGATTATCACCAAGGAAGTGA 693
Db 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
QY 694 ACATTTCTGTTTCAAGCCGAGTAGTGGAAATTTATCAGCTAAAGATGGAGGCCACAGA 753
Db 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
QY 754 TTTAGCAACACTTTTAAGCCCAATCTAGTGGAACTAGTTGTCCCAACATACGATG 813
Db 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
QY 814 GATCCAAATTTATGCCCCGCTGATCCAGACTGGATTGCTTTTATACATGACCAAGTATT 873
Db 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
QY 874 TGGATATCTAACATCGTAAACAGAGAGAAGAGACTCACTTATGTGCACATGAGCTA 933
Db 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu 240
QY 934 GCCAATCATGGAAGAAGATGCCAGATCAGCTGAGTGCCTACCTTTGTCTCCCAAGAGAA 993
Db 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu 260
QY 994 TTTGATAGATATTTCTGGCTATTGTTGGTGTCCAAAGCTGAAACAACTCCCAAGTGTGT 1053
Db 261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrSer----- 276
QY 1054 AAAATCTTTAGAAATCTATATGAAGAAAATGATGAATCTGAGTGGAAATTTATCATGTT 1113
Db 276 ----- 276
QY 1114 ACATCCCCTATTGTAACAAAGGAGGCGAGATTCAATCCGTTATCTCTAAACAGGTACA 1173
Db 277 -----Thr 277
QY 1174 GCMAATCTAAAGTCACCTTTAAGATGTCAGAAATATGATTGATCTGCAAGAGGATC 1233
Db 278 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 297
QY 1234 ATGATGTCATAGATAAGAACTAATTCACCTTTTGAGATTCTTATTTGAGGAGTTGAA 1293
Db 298 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 317
QY 1294 TATATTGCGAGAGCTGGATGAGACTCTCTGAGGGAATAATATGCTGGTCCATCTCTAGAT 1353
Db 318 TyrIleAlaArgAlaGlyTyrThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 337
QY 1354 CGCTCCAGACTCGCCTACAGATAGTGTGTGATCTCACTGAAATTTATTTATCCAGTAGAA 1413
Db 338 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 357
QY 1414 GATGATGTTTATGGAAGGCGAGACTCAATTGAGTCAGTGGCTGATCTGTGACGCCACTA 1473
Db 358 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 377
QY 1474 ATTATCTTGAAGAAACACACATCTGGATAAATATCCATGACATCTTTTCATGTTTTT 1533
Db 378 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 397

QY 1534 CCCCAAAGTCACGAAGAGAAATTCAGTTTATTTTTCCTCTGAATCCAAACAGGTTTC 1593
Db 398 ProGlnSerHisGluGluGluIleGluPheAlaSerGluCysLysThrGlyPhe 417
QY 1594 CGTCATTTATACAAAATACATCTATTTTAAAGGAAAGCAAAATATAAAGCATCCAGTGGT 1653
Db 418 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerGlyGly 437
QY 1654 GGCTGCTGCTCCCAAGTGATTTCAAGTGTCTCTATCAAGAGGAGATAGCAATTTACCACT 1713
Db 438 GlyLeuProAlaProSerAspPheLysCysProIleLysGluIleAlaIleThrSer 457
QY 1714 GGTGAATGGGAAGTCTTTGGCGGCATGGATCTAATATCCAAAGTTGATGAAGTCAGAAG 1773
Db 458 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 477
QY 1774 CTGCTATATTTTGAAGGACCAAGACTCCCTCTTTAGAGCATCACCTGTACGTAGTCAGT 1833
Db 478 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSer 497
QY 1834 TAGCTAAATCTCGGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACATTTCTTGCTGC 1893
Db 498 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 517
QY 1894 ATCAGTCAGCACTGTGACTTCTTTATAAGTAAAGTATAGTAAACAGAGAATCCACACTGT 1953
Db 518 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 537
QY 1954 GTGTCCCTTTAAGCTATCAAGTCTGAGTCCAGAGTCAAGTCCCACTTCGAAAACAAAGGAATTT 2013
Db 538 ValSerLeuTyrLysLeuSerSerProGluAspAspProLeuAla---AsnLysGluPhe 556
QY 2014 TGGGCCACCATTTTGGATTCAGCAGGTCTCTCTGACTATCTCTCTCCAGAAATTTTC 2073
Db 557 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 576
QY 2074 TCTTTTGAAGTACTACTGGATTTACATTTAGTGGATGCTCTCAAGCCTCATGATCTA 2133
Db 577 SerPheGluSerThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 596
QY 2134 CAGCTCGGAAGAAATATCTACTGCTGTCTCATATATGTTGCTCTCAGGTGCAAGTTG 2193
Db 597 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGln----- 613
QY 2194 GTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTTAGCCTCTCTAGGT 2253
Db 613 ----- 613
QY 2254 TATGTGTTGTAGTAGACAAACAGGGGATCCTGTCCACGAGGGCTTAAATTTGAAAGC 2313
Db 613 ----- 613
QY 2314 GCCTTTAAATATAAAATGGTCAAATAGAAATTGACGATCAGGTGGAAGGACTCCAATAT 2373
Db 613 ----- 613
QY 2374 CTAGCTTCTCGATATGATTTTCATTGACTTAGATCGTGTGGGACCCACGGCTGCTCTAT 2433
Db 613 ----- 613
QY 2434 GGAGGATACCTCTCCCTGATGCATTAATGCAGAGGTGAGATATCTTCAGGGTTGCTATT 2493
Db 614 -----ValaIle 616
QY 2494 GCTGGGGCCCCAGTCACTCTGTGATCTCTTATGATACAGATACCGAACGTTATATG 2553
Db 617 AlaGlyProProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet 636
QY 2554 GGTCACTCCGACAGAAATCAACAGGCTATTAATTAGGATCTGTGGCCATCGCAAGCGAA 2613
Db 637 GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 656

QY 2614 AAGTTCCCTCTGACCAAAATCGTTTACTGCTCTTACATGGTTTCTCGATGAGAAATGTC 2673
DB 657 LysPheProSerGluProAsnArgLeuLeuLeuLeuLeuHisGlyPheLeuAspGluAsnVal 676
QY 2674 CATTTTGCACATACCACTATATTAATCTAGTCTTTTATAGTCAGGCTGGAAAGCCATATGAT 2733
DB 677 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 696
QY 2734 TTACAGATCTATCTCTAGGAGACACAGCATTAAGATTCTCCTGAATCGGAGAACATTAT 2793
DB 697 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 716
QY 2794 GAACCTGCATCTTTTGCACCTACCTTCAAGAAACCTTGGATCAGCTATTGCTGCTCTAAAA 2853
DB 717 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeuLys 736
QY 2854 GTGATA 2859
DB 737 Vallie 738
RESULT 14
ABB97362
ID ABB97362 standard; protein; 724 AA.
XX AC ABB97362;
XX DT 27-JUN-2002 (first entry)
XX Novel human protein SEQ ID NO: 630.
DE Human; antianaemic; vulnary; antiinflammatory; immunomodulator;
KW antinfertility; cerebroprotective; cyrostatic; rheumatic; gene therapy;
KW neuroprotective; antiparkinsonian; protein therapy; ESI;
KW expressed sequence tag.
XX Homo sapiens.
OS WO200222660-A2.
XX PN 21-MAR-2002.
XX PD 10-SEP-2001; 2001WO-US026015.
XX PF 11-SEP-2000; 2000US-00659671.
XX PA (HYSE-) HYSEQ INC.
PI Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;
XX WPI; 2002-292408/33.
DR N-PSDB; ABB32548.
XX An isolated polynucleotide for treating diseases associated with its
PT encoded polypeptide such as cancer and multiple sclerosis.
XX Example 2; SEQ ID NO 630; 509pp; English.
XX The present invention provides the protein and coding sequences of 444
CC novel human proteins. These were isolated from expressed sequences tags
CC (ESTs). They can be used to stimulate cell growth, to regulate
CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
CC e.g. in burn treatment, to regulate the immune system e.g. to treat
CC multiple sclerosis, to regulate activin or inhibit e.g. to treat
CC infertility, to regulate haemostasis or thrombolysis e.g. to treat stroke
CC and cancer, to screen for drugs, to treat inflammatory conditions e.g.
CC rheumatoid arthritis, and to treat nervous system disorders e.g.
CC Parkinson's disease. The present sequence is a protein of the invention
SQ Sequence 724 AA;

Alignment Scores:
Pred. No.: 0 Length: 724

Score: 3771.00 Matches: 724
Percent Similarity: 82.1% Conservative: 0
Best Local Similarity: 82.1% Mismatches: 0
Query Match: 67.9% Indels: 158
DB: 5 Gaps: 2
US-10-825-632-2 (1-3120) x ABB97362 (1-724)
QY 214 ATGGCAGCAGCAATCGAAACAGAACAGCTGGGTGTTGAGATATTTGAAACTGCGGACTGT 273
DB 1 MetAlaAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
QY 274 GAGGAGAAATATGAATATCAGAGTCGCGCTTAATATGGAGCCCTTTTATGTTGAGCGGTAT 333
DB 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
QY 334 TCCTGGAGTCAGCTTAAAAAGCTGCTGCCGATACACAGAAATATCATGCTACATCATG 393
DB 41 SerTyrPserGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
QY 394 GCTAAGGCCACCATGATTTTTCATGTTGTGAAGAGGAATGATCCAGATGGACCTCATTC 453
DB 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
QY 454 GACAGATCTATTACCTTGCCTGCTGCTGAGAACAGAGAAATACACCTGTTTATTCT 513
DB 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
QY 514 GAAATCCCAAACTATCAATAGACGACAGCTCTTAATGCTCTCTGGAAGCCCTCTTTTG 573
DB 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTyrLysProLeuLeu 120
QY 574 GATCTTTTTCAGCAACACATGGAATGATTTCTCGAAGAAGAAGCACTATTAAGA 633
DB 121 AspLeuPheGln----- 124
QY 634 GAAAGAAAACGCATTGGAAACAGTCGGAATTCCTTACGATTTATCACAAGGAAGTGA 693
DB 124 ----- 124
QY 694 ACATTTCTGTTCAAGCCGCTAGTGAATTTATCAGCTAAAGATGGAGGCCCAAGGA 753
DB 124 ----- 124
QY 754 TTTACGCAACACCTTTAAGGCCCAATCTAGTGGAAACTAGTTGTGCCAACATACGATG 813
DB 125 -----GlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 142
QY 814 GATCCAAAATATGCCCCGCTGATCCAGCTGGATGCTTTTATACATAGCAACGATATT 873
DB 143 AspProLysLeuCysProAlaAspProAspTyrIleAlaPheIleHisSerAsnAspIle 162
QY 874 TGATATCTAATCTGTAACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 933
DB 163 TrpIleSerAsnIleValThrArgGluGluArgArgLeuThrTyrValIleAsnGluLeu 182
QY 934 GCCAACATGGAAGAAGATCCAGATCAGCTGAGGTGCTGCTACCTTTGTTCTCCAAAGAGAA 993
DB 183 AlaAsnMetGluGlnAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu 202
QY 994 TTTGATAGATATTTCTGCTATTGGTGGTGTCCAAAGCTGAAACAACTCCCGAGTGGTGT 1053
DB 203 PheAspArgTyrSerGlyTyrTrpTyrCysProLysAlaGluThrThrProSerGlyGly 222
QY 1054 AAAATCTTGAATCTTATATGAGAAATGATGATCTGAGGTGGAATATTTTCATCTT 1113
DB 223 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 242
QY 1114 ACATCCCTGATGTTGAAACAAGAGGCGAGATTCATTCCTGCTATCTTCTTCTTCTTCTTCT 1173
DB 243 ThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThr 262
QY 1174 GCMAATCCTAAAGTCATCTTTTAAGATGTCCAGAAATATGATGATGCTGCTGAAGGAGGATC 1233

Db	263	AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle	282		599	-----	599
QY	1234	ATAGATGTCATAGATAAGAACTAAATTCACCTTTTGAGATTTCTATTTGAAGGAGTTGAA	1293		2314	GCCTTTAAATATAAATGGGTCAAAATAGAAATTGACGATCAGGTGGAGGACTCCAATAT	2373
Db	283	IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu	302		599	-----	599
QY	1294	TATATTGCCAGAGCTCGATGGAGCTCTCTGAGGGAATAATGCTTGGTCCATCTCTACTAGAT	1353		2374	CTAGCTTCTCGATATGATTTTCATTGACTTAGATCGTGTGGGCATCCACGGCTGGTCTAT	2433
Db	303	TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp	322		599	-----	599
QY	1354	CGCTCCAGACTCGCCTACAGATAGTGTGATCTCACCTGAATATTATATCCAGTAGAA	1413		2434	GGAGGATACCTCTCCCTGATGGCATTAATGCAGAGGTCCAGATATCTTCCAGGTTGCTATT	2493
Db	323	ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu	342		600	-----	602
QY	1414	GATGATGTTATGGAAGGACAGAGACTCATTGAGTCAGTCGCTGATTCGTGACGCCACTA	1473		2494	GCTGGGGCCCCAGTCACCTCTGTGGATCTTCTATGATACAGATACACGGAACGTTATATG	2553
Db	343	AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu	362		603	AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet	622
QY	1474	ATTATCTATGAAGAAACAACAGACATCTGGATAAATATCCATGACATCTTTCATGTTTTT	1533		2554	GGTCACCTCTGACAGAAATGAACAGGGCTATTACTTTAGGATCTGTGGCCATGCCAAGCAGAA	2613
Db	363	IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe	382		623	GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu	642
QY	1534	CCCCAAAGTCACGAGGAAATTTGAGTTTATTTTGGCTCTGATGCAAAACAGGTTTC	1593		2614	AAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTACATGTTTCTCTGGATGAGAATGTC	2673
Db	383	ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe	402		643	LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal	662
QY	1594	CGTCATTTATACAAATTTACATCTATTTTAAAGGAAGCAATATATAAAGATCCAGTGGT	1653		2674	CATTTTGACATACACAGATATATTACTGAGTTTTTTTACTGAGGGCTGGAAAGCCATATGAT	2733
Db	403	ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly	422		663	HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp	682
QY	1654	GGCTGCTGCTCCAAAGTGATTTCAAGTGTCTCTATCAAGAGAGGATAGCAATTTACAGT	1713		2734	TTACAGATCTATCTCTCAGGAGAGACACAGCATAGAGTTCTCTGAATCGGAGAACATTAT	2793
Db	423	GlyLeuProAlaProSerAspPheLysCysProIleLysGluIleAlaIleThrSer	442		683	LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr	702
QY	1714	GGTGAATGGGAAGTCTTGGCCGCGATGATCTAATATCCAAGTTGATGAAGTCAGAAGG	1773		2794	GAACTGCATCTTTTGCACCTACCTTCAAGAAAACCTTCGATCAGTATTCGCTCTTAAAA	2853
Db	443	GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg	462		703	GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeuLys	722
QY	1774	CTGGTATATTTTGAAGGCACAAAGACTCCCCCTTTTAGAGCATCACCTGTACGTAGTCAGT	1833		2854	GTGATA	2859
Db	463	LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSer	482		723	ValIle	724
QY	1834	TACGTAATCTCGAGAGGTGACAAAGGCTGACTGACCGGTGCTACTCACATTTCTTGCTGC	1893		RESULT 15		
Db	483	TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys	502		ID	ABG61594	standard; protein; 690 AA.
QY	1894	ATCAGTCACAGCTGACTCTCTTTATAGTAGTAGTATGTAACCCAGAGATCCACACTGT	1953		XX	ABG61594;	
Db	503	IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys	522		AC	ABG61594;	
QY	1954	GTGTCCTTTACAGCTATCAAGTCTCTGAGATGACCCCACTTGCAAAACAAAGGAATTT	2013		XX	12-AUG-2002 (first entry)	
Db	523	ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe	542		DT	Human DRP-1 splice variant #1.	
QY	2014	TGGGCCACATTTTGGATTGAGAGGTCTCTCTCTGACTATATCTCTCCAGAAATTTTC	2073		DE	Human DRP-1 splice variant #1.	
Db	543	TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe	562		XX	Human; serine protease; dipeptidyl peptidase IV-related protein; DPPP;	
QY	2074	TCTTTTGAAGTACTACTCGATTTATCATTTGATGGGATGCTCTACAAGCCTCATGATCTA	2133		KW	DPP4V; infection; human immunodeficiency virus; HIV-1; HIV-2; pain;	
Db	563	SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu	582		KW	diabetes; infertility; obesity; anorexia; Parkinson's disease; stroke;	
QY	2134	CAGCTCGGAAGAAATATCTCTACTGTGCTGTTTCATATATGTGGTGGTCTCAGGTGCAGTTG	2193		KW	heart failure; hypertension; urinary retention; osteoporosis; cancer;	
Db	583	GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGln	599	-----	KW	ulcer; allergy; cancer; psychotic disorder; neurological disorder;	
QY	2194	GTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATATCCCTAGCCTCTCTAGGT	2253		KW	dyskinesia; reproductive disorder; inflammatory disorder;	
Db	599	-----	599	-----	XX	metabolic disorder.	
QY	2254	TATGTGTTGTAGTAGACAAACAGGGGATCCTGTCACCGAGGGCTTAAATTTGAAGGC	2313		OS	Homo sapiens.	

XX WPI; 2002-444178/47.
DR N-PSDB; ABR83325.
XX New dipeptidyl peptidase IV-related proteins and nucleic acids encoding
PT the proteins, useful for treating e.g. fungal, bacterial, protozoan and
PT viral infections, cancers, allergies, neurological disorders, or pain.
XX Disclosure; Page 59-61; 113pp; English.
PS The present invention relates to the isolation of novel human serine
PS proteases referred to as dipeptidyl peptidase IV (DPP-IV)-related proteins
CC (DPP-IV). The dipeptidyl peptidase IV-related proteins (DPP-IV) and nucleic
CC acids encoding them are useful for treating infections such as fungal,
CC bacterial, protozoan and viral infections, particularly infections caused
CC by human immunodeficiency virus (HIV-1 or HIV-2), pain, diabetes,
CC precocious puberty, infertility, obesity, anorexia, bulimia, Parkinson's
CC disease, acute heart failure, hypotension, hypertension, urinary
CC retention, osteoporosis, angina pectoris, stroke, ulcers, asthma,
CC allergies, cancers, migraine, vomiting, psychotic and neurological
CC disorders (e.g. anxiety, dementia, or schizophrenia), and dyskinesias.
CC These may also be used in discovering therapeutic agents for the
CC treatment of reproductive, inflammatory and metabolic disorders. ABG61591
CC -ABG61612 represent human DPP-IV proteins
XX
XX Sequence 690 AA;

Alignment Scores:
Pred. No.: 0 Length: 690
Score: 3607.50 Matches: 690
Percent Similarity: 87.1% Conservative: 0
Best Local Similarity: 87.1% Mismatches: 0
Query Match: 65.0% Indels: 102
DB: 5 Gaps: 1

US-10-825-632-2 (1-3120) x ABG61594 (1-690)

QY 214 ATGGCAGCAGCAATGGAACAGACAGCTGGGTGTGAGATATTGAACTGCGGACTGT 273
DB 1 MetAlaAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
QY 274 GAGGAGATATTGAATCAAGATCGGCTAAATGGAGCGCTTTTATGTTGAGCGGTAT 333
DB 21 GluGluAsnIleGluSerGlnAspArgProlysLeuGluProPheTyrValGluArgTyr 40
QY 334 TCTGAGTACGCTTAAAGCTGCTTCCGATACCAGAAATATCATGCTCATGATG 393
DB 41 SerTrpSerGlnLeuLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
QY 394 GCTAAGCCACCATGATTTCATGTTGTGAAGAGGAATGATCCAGATGACCTCATTTCA 453
DB 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
QY 454 GACAGAACTATTACCTTCCCATGTCTGGTGAAGACAGAGAAATACACTGTTTATCT 513
DB 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
QY 514 GAAATCCAAAATCATATAGACAGCAGTCTTAATGCTCTCTTGGAGCCCTCTTTG 573
DB 101 GluIleProLysThrIleAsnArgAlaValLeuMetLeuSerTrpLysProLeuLeu 120
QY 574 GATCTTTTTCAGCAGCAGTACGATGATGATGATTTCTCGAGAGGAAGAACTATTAGA 633
DB 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuArg 140
QY 634 GAAAGAAAACGCAATCGAAGCTGCGAATGCTTCTTACGATTATCATCCAGGAGTGA 693
DB 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
QY 694 ACATTTCTGTTTCAAGCCGGTAGTGAATTTATCATACGTAAAGATGAGGGCCACAGGA 753
DB 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180

QY 754 TTTACGCAACACCTTTAAGGCCCAATCTAGTGGAAACTAGTGTGCCAACATACGATG 813
DB 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
QY 814 GATCCAAAATATTGCCCCCTCATCCAGACTGGATTGCTTTTATACATAGCAAGATATT 873
DB 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
QY 874 TGGATATCTAAACATCTGTAAACCCAGAGAGAAAGAGAGACTCATTATGTGCACATGAGCTA 933
DB 221 TrpIleSerAsnIleValThrArgGluArgLeuThrTyrValHisAsnGluLeu 240
QY 934 GCCAAATCGAAGAGATCCAGATCAGCTGAGTGCCTACTCTTGTCTCCAAAGAGAA 993
DB 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu 260
QY 994 TTTGATAGATATTCTCGCTATTGGTGTCTCCAAAGCTGAAACAACTCCAGTGGTGGT 1053
DB 261 PheAspArgTyrSerGlyTyrTrpCysProLysAlaGluThrThrProSerGlyGly 280
QY 1054 AAAATTTCTAGAAATCTATATGAAGAAATGATGAATCTGAGGTGGAAATTTATTCATGTT 1113
DB 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300
QY 1114 ACATCCCTATTGTTGAAACAGAGGGGCGAGATTCTATCCGTATCTCTAAACAGGTACA 1173
DB 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320
QY 1174 GCAATCTCTAAAGTCACTTTTAAGATGTCCAGAAATAATGATTGATGCTGAAGAGGATC 1233
DB 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
QY 1234 ATAGATGTCTATAGATAAGGAACATAATCAACTTTGATGATTCATTTTGAAGGAGTTGAA 1293
DB 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
QY 1294 TATATGCCAGACTGGATGGACTCTGAGGAGAAATATGCTGTGCTCCATCTCTACTAGAT 1353
DB 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
QY 1354 CGCTCCAGACTCGCTACAGATAGTGTGTGATCTCACCTGAAATATTTATCCAGTAGAA 1413
DB 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
QY 1414 GATGATGTATGAAAAGGAGAGACTCATTTAGTCTGAGTCCCTGATTTCTGTGACGCCACTA 1473
DB 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
QY 1474 ATTATCTATGAAGAAACACAGACATCTCGATTAATATCCATGACATCTTTCATGTTTTT 1533
DB 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
QY 1534 CCCCAGAGTCCAGAGGAGAAATGAGTTTATTTTGTGCTGAAATGCAAAACAGAGTTTC 1593
DB 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
QY 1594 CGTCATTTATACAAAATTAATCATCTATTTTAAAGGAAAGCAAAATATAACGATCCAGTGGT 1653
DB 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
QY 1654 GGGTGTCTCTCAGAGTGAATTCAGTGTCTCTATCAAGAGAGATAGCATTAATACCATG 1713
DB 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 500
QY 1714 GGTCAATGGGAGTCTTCTGCGCGCATGGATCTTAATATCCAAGTTGATGAAGTCAGAGG 1773
DB 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
QY 1774 CTGTGTATTTTGAAGGCCAACAAAGACTCCCTCTTAGAGCATCACCTGTACGTAGTCAGT 1833
DB 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisIleLeuTyrValValSer 540
QY 1834 TACGTAATCTCTGAGAGGTTGACAGGCTGACTGACCGGTGGCTACTCAATCTCTGCTGTC 1893

```
Db      541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
Qy      1894 ATCAGTCAGCCTGTGACTTCTTTATAGTAAGTAAGTAAACACAGAGAATCCACACTGT 1953
Db      561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
Qy      1954 GTGTCCTCTTTACAAGCTATCAAGTCCTGAAGATGACCCAACTTGCAAAACAAAAGGAATTT 2013
Db      581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
Qy      2014 TGGGCCACCATTTTGGATTTCAGCAGGTCCTCTTCTGACTATACTCTCCAGAAAATTTTC 2073
Db      601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 620
Qy      2074 TCTTTTGAAGTACTACTGGATTACATTTGATGGATGCTCTACAAAGCTCATGATCTA 2133
Db      621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
Qy      2134 CAGCCTGGAAGAATATCCTACTGTCTGTCTATATATGTTGGTCTCTCAGGTGCAGTTG 2193
Db      641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGly----- 655
Qy      2194 GTGAATAATCGTTTAAAGGAGTCAAGTAATTCGGCTTGAATACCCCTAGCCTCTCTAGGT 2253
Db      655 ----- 655
Qy      2254 TATGTGGTTGTAGTGATAGACAACAGGGGATCCTGTCAACGAGGGCTTAAATTTGAAGGC 2313
Db      655 ----- 655
Qy      2314 GCCTTTAAATATATAAATGGGTCAAAATAGAAATTGACGATCAGGTGGAAGGACTCCANTAT 2373
Db      655 ----- 655
Qy      2374 CTAGCTTCTCGATATGATTTCAATTGACTTAGATCGTGTGGGCATCCACGGCTGGTCCCTAT 2433
Db      655 ----- 655
Qy      2434 GGAGGATACCTCTCCCTGATGGCATTAAATGCAGAGGTACAGATATCTTCAGGGTTGCTATT 2493
Db      656 -----Arg-LeuLeuLe 659
Qy      2494 GCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACGTTATATG 2553
Db      659 uLeuGlyProGlnSerLeuCysGlySerSerMetIleGlnAspThrArgAsnValIleTr 679
Qy      2554 GGTCAACCTGACCAAGTAACAGGGCTATTACT 2587
Db      679 pValThrLeuThrArgMetAsnArgAlaIleThr 690
```

Search completed: April 14, 2006, 11:08:14
Job time : 557.5 secs

Qy	430	AATGATCCAGATGACCTCATTACAGACAGAACTATTACCTTGGCATCTCTGGTGAGAAC	489	Qy	1288	-----GTTGAATATATTGCCAGAGCTGGATGAGTCCCTGAGGAGAAATATGCTTGGTCC	1341
Db	73	MetArgAlaGlu-----LysAspArgLeuAsnMetTyrAlaIleSerSerValPro	89	Db	388	ProPheTyrGluTyrIleThrArgAlaGlyPhePheSerAspGlyThrThrValTyrVal	407
Qy	490	AGAGAAATACA-----CTGTTTATTCTGAAATTCCTCC-----AAAACCTATCAATAGAGCA	540	Qy	1342	ATCCTACATAGATGCTCCAGACTCCGCTACAGATAGTGTGTATC-----	1386
Db	90	GlyThrAsnThrGlnSerIlePheSerValThrIleProLeuGluLeuValGluLysAla	109	Db	408	GlnValMetSerArgAspGlnAlaGlnCysSerLeuLeuLeuIleProTyrThrAspPhe	427
Qy	541	GCATCTTAATGCTCTCTTGGAGCCTCTTTGGATCTTTTTCAGGCAACACTCGACTAT	600	Qy	1387	-----TACCTGAA	1395
Db	110	GlnValAlaAspArgLys-----PheGluLeuLysLeuLysSer	122	Db	428	LeuLeuProGluGluLeuGlyGlySerIleLysGluAspAsnLeuLysLeuSerThrAsp	447
Qy	601	GGATGTAT-----TCTCGAGAA	618	Qy	1396	TTATTATCCAGTAGAGATGATGTTATGGAAGCGAGAGACTCATTTAGTCAGTGCCT	1455
Db	123	GlyTyrAsnValAspSerTyrIleArgMetSerCysArgLysThrProProSerAlaGlu	142	Db	448	LeuAsnMetGlyValTyrAspAsp-----LysSerHisGluGluThrMetGluLysProPro	466
Qy	619	GAAGAACTATTAGAGAGAGAAACCAATGGACAGTC-----GGAATGCTCTTACCAT	675	Qy	1456	GATTCTGTGAGC-----CCACTAATATCTATGAAGAAACAAACAGACATCTCGATAAAT	1509
Db	143	PheThrLeuGlnCysGluArgGlnArgSerGlnValValThrGlyIleSerAspTyrGlu	162	Db	467	ArgGlyLysLeuArgGlyThrValGlnIleHisLysAlaArgAsnAspTyrTyrIleAsn	486
Qy	676	TATCACCAAGGA-----AGTGGAACTATTTCTGTTTCAAGCCGT-----	714	Qy	1510	ATCCATGATCATCTTCATGTTTTCCTCCCAAGTCCAGAGAGGAA-----ATTGAGTTT	1563
Db	163	IleArgAsnGlyLysMetIleLeuMetAlaGlyAspGlnLeuPheArgTyrAsnProLeu	182	Db	487	ThrHisAsnAlaIleTyrProLeuLysIleThrAspGluGluHisProMetTyrGluPhe	506
Qy	714	-----	714	Qy	1564	ATTGTTGCTCTGATGCAAAACAGGTTTCGTCATTATACAAATATCATCTATTTTA	1623
Db	183	AsnGluAlaLeuAlaAlaIleProIleAlaValProAspAspGlnSerSerThrGluPro	202	Db	507	IleTyrCysLeuGluLysProAsnGly-----SerCysLeuAlaLeuIleSerAlaGluLeu	525
Qy	715	-----AGTGGAAATTTATCAGCTAAAGATGAGGGCCA	747	Qy	1624	AAGGAAGCAATATAACGATCCAGTGGTGGCTGCTCTCCAGTGATTTCAGTGT	1683
Db	203	MetAspIleSerGluGlySerIleThrSerGlyThrLysGlyCysSerAsnGluAlaPro	222	Db	526	AspGlnAsnGlyTyr-----Cys	531
Qy	748	CAAGGA-----TTTACGCAACAACTTTAAGGCCCAATCTAGTGGAACT	792	Qy	1684	CCTATCAAGAGAGATAGCAATATACAGTGTGTAATGGGAGTTCTTGGCCGCATGGA	1743
Db	223	GlnSerSerThrValProValThrArgIleProIleLysLysProThrThrSerThr	242	Db	532	ArgHisThrGluGluLysLeuLeuMetAlaGluAsnPheSerIle-----AsnLysSer	549
Qy	793	AGTTCTGCC-----AACATACGATGATCCAAATATGATGCCCC	831	Qy	1744	TCTAATATCAAGTTGATGAAAGGCTGATATATTTTGGAGCCACCAAGACTCC	1803
Db	243	GluLysProAlaThrAlaProProThrAsnAsnPheValSerSerAlaLysValCysPro	262	Db	550	MetGlyIleValValAspGluValArgGluLeuValTyrTyrValAlaAsnGluSerHis	569
Qy	832	GCTGATCCAGACTGGATGCTTTATACATAGCAACGATATTTGGATATCTAACATCGTA	891	Qy	1804	CCTTTAGACATCACCCTGTACGTAGTACGTAAATCTGAGAGGTGACAGGCTG	1863
Db	263	AlaAspSerSerLeuLeuAlaTyrValLeuAsnLysGlnValTyrIle-----	278	Db	570	ProThrGluTyrAsnIle-----CysValSerHisTyrArgThrGlyGlnHisAlaGluLeu	588
Qy	892	ACCAGAGAAAGAGAGACTCATTATGTGCACATGAGCTAGCCACATGGAAGAAGAT	951	Qy	1864	ACTGACCGTGGCTACTCACAATCTTGTCTGCATCAGTCAGCACTGTCTTTTATAAGT	1923
Db	279	-----GluLysAsnGlyLysIleIleHisArgThrSerSerAsnSerLysHisIle	295	Db	589	ThrGluSerGlyIle-----Cys-----PheLysSer	597
Qy	952	GCAGATCAGCTGGAGTGCCTTCTCTCCAGAGAAATTTGATAGATATCTGCG	1011	Qy	1924	AAGTATAGTAACCAAG-----AATCCACACTGTGTCTCTTTCACAGCTA	1971
Db	296	ThrAsn-----GlyValProSerTyrIleValGlnGluLeuGluArgPheGluGly	313	Db	598	GluArgAlaAsnGlyLysLeuAlaLeuAspLeuAspHisGlyPheAlaCysTyrMetThr	617
Qy	1012	TATGTGTGTGTCAAAAGCTGAAACAACTCCAGTGGTGTAAATTTCTAGAAATCTA	1071	Qy	1972	TCA-----AGTCTGGAAGATGACCAACTTGCAAAACAAAGGAATTT---TGG-----	2016
Db	314	IleTyrTrp-----SerGluSerLysThr-----ArgLeuLeu	324	Db	618	SerValGlySerProAlaGlu-----CysArgPheTyrSerPheArgTyrLysGlu	634
Qy	1072	TATGAAGAAATGATGAATCTGAGGTGGAATTAATTCATGTT-----	1113	Qy	2017	-----GCCACATTTTGGATTTCAGAGGTCTCT---	2043
Db	325	TyrGluHisValAsnGluLysValAlaGluSerGlnPheGlyValAsnGlyAspPro	344	Db	635	AsnGluValLeuProSerThrValTyrAlaAlaAsnIleThrValSerGlyHisProGly	654
Qy	1114	---ACATCCCTCTATGTTGGAAACAGGAGGCGAGATTCATTCCGTTATCTCTTAAACAGGT	1170	Qy	2044	CTTCTCTGAC-----TATACCTCTCAGAAATTTCTCTTTTGAAGT---ACTACTGGA	2094
Db	345	ProValAlaProMet-----LysTyrProArgAlaGly	355	Db	655	GlnProAspLeuHisPheAspSerProGluMetIleGluPheGlnSerLysLysThrGly	674
Qy	1171	ACAGCAAAATCTTAAAGTCACTTTTAAGATGTGAGAAATATGATGCTGAAGGAAGG	1230	Qy	2095	TTTACATTTGATGGATGCTCTACAGCTCATGATCTACAGCTGGAAGAAATATCCT	2154
Db	356	ThrLysAsnAlaTyrSerThrLeuArgMet-----ValIleLeuGluAsnGlyLys	372	Db	675	LeuMetHisTyrAlaMetIleLeuArgProSerAsnPheAspProTyrLysLysTyrPro	694
Qy	1231	ATCATAGATGTCATAGATAAGCACTAATTCACCTTTTGAAGATTCATTATTGAAGA---	1287	Qy	2155	ACTGTGCTGTTCATATATGCTGCTCTCAGTGGCAATGCTGCTGAATATCGGTTTAAAGGA	2214
Db	373	AlaTyrAspValProLeuLysAsp-----GluValIleTyrLysHisCys	387	Db	695	ValPheHisTyrValTyrGlyGlyProGlyIleGlnIleValHisAsnAspPheSerTrp	714
Qy				Qy	2215	GTCAAGTATTTCCGCTTGAATACCCCTAGCCCTCTCTAGGTTATGCTGTTGTAGTAGATAGAC	2274

Db	715	IleGlnTyrIleArg-----PheCysArgLeuGlyTyrValValPheIleAsp	731	QY	436	CCAGATGGACCTCATTCCAGACAGAACTCTATTACCTTGCATGTCTGTGTGAGAACAGAGAA	495
QY	2275	AACAGGGGATCCTGTCCACGAGGGCTTAAATTTGAAGCGCCTTTAAATATAAAATGGGT	2334	Db	47	ProAspGly-----LysArgValThrTyrLeuLysGlyLysProGlu-----	60
Db	732	AsnArgGlySerAlaHisArgGlyIleGluPheGluArgHisIleHisLysLysMetGly	751	QY	496	AATACACTGTTTTATTCTGAAATTCCTCAAACTATCAATAGAGCAGCAGCTTTAATGCTC	555
QY	2335	CAAAATGAATTCAGCATCAGGTGGAAGACTCCCAATATCTAGCTTCTCGATAT- --GAT	2391	Db	61	-----AlaAlaAsnIleGlnAsp	66
Db	752	ThrValGluValGluAspGlnValGluGlyLeuGlnMetLeuAlaGluArgThrGlyGly	771	QY	556	TCITGG-----AAGCCTCTTTTGGATCTTTTTCAGGCCAACACTG	594
QY	2392	TTCAATGACTAGATCGTGTGGCATCCACGGCTGCTCTATGAGGATACCTCTCCCTG	2451	Db	67	LeuTrpAlaAlaAspValLysGlyGluProTyrArgLeuIleAspSerAlaAlaLeu	86
Db	772	PheMetAspMetSerArgValValHisGlyTyrSerTyrGlyTyrMetAlaLeu	791	QY	595	GACTATGGA-----ATGTATTCTCAGAGAAGAACTATTAAAGAGAAAGAAACGCATT	648
QY	2452	ATGGCATTATGCAGAGTCCAGATATCTTCAGGGTGTCTATTGCTGGGGCCCGCCAGTCACT	2511	Db	87	SerSerGlyAspLysGluLeuSerGluAlaGluLysAlaArgGluArgAlaArgVal	106
Db	792	GlnMetIleAlaLysHisProAsnIleTyrArgAlaAlaIleAlaGlyAlaValSer	811	QY	649	GGAACTCGGAATTCCTTACGATTATCACCAAGGAAGTGGACACATTTCTGTTTCAA	708
QY	2512	CTGTGGATCTTCTATGATACAGATACAGGAACGTTATATGGCTCACCTCGACACAGAAT	2571	Db	107	SerAlaArgGlyIleValGluTyrSerTrpAspArgGlnGlyArgPheIleLeuValPro	126
Db	812	AspTrpArgLeuTyrAspThrAlaTyrThrGluArgTyrMetGlyTyrPro- --LeuGlu	830	QY	709	GCCGTAGTGGATTTAT-----CACGTAAAGATGCGAGGCCACCAAGGATTTACGCCAA	762
QY	2572	GAACAGGCTTATTAGGATCTGTGGCCATGCAAGCAGAGAAAAGTTCCCTCTGAACCA	2631	Db	127	LeuAspGlyAspLeuTyrLeuAspAlaValAlaAspGlyLysIleThrArgLeuThrGlu	146
Db	831	GluHisValTyrGlyAlaSerSerIleThrGlyLeuValGluLysLeuProAspGluPro	850	QY	763	CAACCTTTAAGGCCCAATCTAGTGGAACTAGTTCTCCCAACATACGATCGATCCAAAA	822
QY	2632	AATCGTTTACTGCTTACATGCTTTCTCGATGAGATGCAATTTGTCACATACACGT	2691	Db	147	-----ThrProGlyAspGluValAspAlaLys	155
Db	851	AsnArgLeuMetLeuValHisGlyLeuMetAspGluAsnValHisPheAlaHisLeuThr	870	QY	823	TTATGCCCGCTGATCCAGACTGGATGCTTTTATACATACCAACGATTTTGGATATCT	882
QY	2692	ATATTACTAGTTTTTACTGAGGGCTGGAAGCCATATGATTACAGATCTATCCTCAG	2751	Db	156	ValSerProLysGly--GlyTyrValSerTyrValArgAspGlnAsnLeuTyrIleLys	174
Db	871	HisLeuValAspGluCysIleLysLysGlyLysTrpHisGluLeuValIlePheProAsn	890	QY	883	ACATCGTAAACAGAGAGAAAGGAGACTCACTATGTGCAATGACGATGACCCAAATG	942
QY	2752	GAGACAGACAGCATAGAGTCTCGAATCGGGAGAACATTTGAACTGTCATCTTTGGCAC	2811	Db	175	ProValAlaGlyGlyAlaGluThrAlaLeuThr-----ThrAsp	187
Db	891	GluArgHisGlyValArgAsnAsnAspAlaSerIleTyrLeuAspAlaArgMetMetTyr	910	QY	943	GAAGAGATGCCAGATCAGCTGAGTCCGCTACTCTTTGTTCTCCAGAGAAATTTGATAGA	1002
QY	2812	TACCTTCAAGAA 2823		Db	188	GlyLysAspAlaLeuSerPheGlyValAlaGluPheIleValGlnGluLeuAspArg	207
Db	911	PheAlaGlnGln 914		QY	1003	TATTCTGGCTTATTGGTGTCTCAAAAGCTGAAACAACTCCACGTCGGTGGTAAAAATCTT	1062
RESULT 2				Db	208	PheThrGlyTyrTrpTrpSerProAspGluSer-----	218
dipeptidyl peptidase IV [imported] - Caulobacter crescentus				QY	1063	AGAATCTATATGAAGAAAATCATGAATCTGAGGTGGAATATTTATCATGTTACATCCCT	1122
C;Species: Caulobacter crescentus				Db	219	ArgIleValTyrThrArgValAspGluSerGlyValAspIleVal-----	233
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004				QY	1123	ATGTTGGAACAAGGAGGCGCAGAT-----TCATTCCGT	1155
R;Niernman, W.C.; Feidblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.				Db	234	-----ProArgAlaAspIleGlyProGlyGlyAlaThrValValAsnGlnArg	249
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon				QY	1156	TATCTTAAACAGGTACACAAATCTTAAAGTCACCTTTTAAAGATGTGAGAAATAATGATT	1215
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.				Db	250	TyrProArgAlaGlyArgProAsnAlaValValAspLeuPheValArgAspLeu-----	267
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001				QY	1216	GATGCTGAAGGAGGATC-----ATAGATGTC-----ATAGATAAGGAACATAATTCAACCT	1266
A;Title: Complete Genome Sequence of Caulobacter crescentus.				Db	268	---AlaSerGlyLysValThrAlaLeuAspLeuGlyAlaAsnLysAspIle-----	283
A;Reference number: A87249; MUID:21173698; PMID:11259647				QY	1267	TTTGAGATCTATTGGAAGAGTGAATATATTTGCGAGAGCTGGATGAGCTCTCGAGGA	1326
A;Accession: A87516				Db	284	-----TyrValAlaArgValAlaTrpSerAlaAspGly	294
A;Status: preliminary				QY	1327	AAATATGCTTGGTCACTCTACTAGATCGTCCCGAGCTCCCTTACAGATAGTGTTCATC	1386
A;Molecule type: DNA				Db	295	LysThrValTyrValGlnArgLeuSerArgAspGlnLysThrLeuAspLeuLeuAlaPhe	314
A;Residues: 1-738 <STO>				QY	1387	TCACCTGAATTTATTTATCCAGTAGAAGATGATGTTATGGAAGGCGACAGAGACTCATTTGAG	1446
A;Cross-references: UNIPROT:Q9A6E0; UNIPARC:UPI00000C7616; GB:AE005673; NID:gl3423647; F				Db	314	-----	314
C;Genetics:				QY	1447	TCAGTGCCTGATTTCTGTGAGCGCCACTA-----ATTATCTATGAAGAAACACAGACATC	1500
A;Gene: CC2154							
Alignment Scores:							
Pred. No.:	5,93e-53	Length:	738				
Score:	842.00	Matches:	235				
Percent Similarity:	43.9%	Conservative:	119				
Best Local Similarity:	29.2%	Mismatches:	282				
Query Match:	15.2%	Indels:	170				
DB:	2	Gaps:	28				

US-10-825-632-2 (1-3120) x A87516 (1-738)

Db 129 GlyGluLeuTyrLeuTyrAspLeuLysGlnGluGlyLysAlaAlaValArgGlnLeu--- 147
Qy 769 TTAAGGCCCAATCTAGTGGAACTAGTTGTGCCAACATACGATGGATCCAAATAATTATGC 828
Db 148 -----ThrHisGlyGluGlyPheAlaThrAspAlaLysLeuSer 160
Qy 829 CCCGCTGATCCAGACTGGATTGCTTTTATACATAGCAAGCATATTTGGATATCTAACATC 888
Db 161 ProlysGly---GlyPheValSerPheIleArgGlyArgAsnLeuTrpValIleAspLeu 179
Qy 889 GTAACACAGAGAAGAGACACTCATTTATGTGCACAACTGCCAGTAGCCCAACATCGAAGAA 948
Db 180 AlaSerGlyArgGlnMetGlnLeuThr-----Ala 189
Qy 949 GATCCAGATCAGCT-----GGAGTCGCTACCTTTGTTCTCCAAGAAGAAATTTGAT 999
Db 190 AspGlySerThrThrIleGlyAsnGlyIleAlaGluPheValAlaAspGluGluWetAsp 209
Qy 1000 AGATATCTGGCTATTGGTGGTGTCCAAAGCTGAAACAACTCCCGAGTGGTGGTAAATTT 1059
Db 210 ArgHisThrGlyTyrTrpAlaProAspAspSerAla----- 222
Qy 1060 CTTAGATTCTATATGAAGAAATGATGAATCTGAGTGC-----GAA 1101
Db 223 -----IleAlaTyrAlaArgIleAspGluSerProValProValGlnLysArgTyrGlu 240
Qy 1102 ATTATTCATGTTACATCCCTATGTTGGAAACAAGGAGGCGCAGATTTCATTCCTGTTATCCT 1161
Db 241 ValTyrAlaAspArgThrAspValIleGluGln-----ArgTyrPro 254
Qy 1162 AAAACAGGTACAGCAATCTAAAGTCACCTTTTAAGATGTGCAGAAATAATGATGATGC 1221
Db 255 AlaAlaGlyAspAlaAsnValGlnValLysLeuGlyValIleSerProAlaGluGlnAla 274
Qy 1222 GAAGGAAGGATCATAGATGTCATAGATAGAGAACTAATTCACCTTTTGAGATTCATTT 1281
Db 275 GlnThrGlnTrpIleAspLeuGlyLysGluGlnAspIle----- 287
Qy 1282 GAAGGAGTTGAATATATGTCAGAGCTGGATGAGTCTCTGAGGGAAATATGTTGTGTC 1341
Db 288 -----TyrLeuAlaArgValAsnTrpArgAspProGlnHisLeuSerPheGln 303
Qy 1342 ATCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCACTCCATGAATATTT 1401
Db 304 ---ArgGlnSerArgAspGlnLysLysLeuAspLeuValGluValThrLeu----- 319
Qy 1402 ATCCAGTAGAAGATGATGTATGGAAGGCGAGAGACTCATTTGAGTCAGTGCCTGATTC 1461
Db 320 -----AlaSerAsnGlnArg----- 325
Qy 1462 GTGAGGCCACTAATTATCTATGAAGAAACAACAGACTCTGGATAAATATCATGACATC 1521
Db 326 -----ValLeuAlaHisGluThrSerProThrTrpValProLeuHisAsnSer 341
Qy 1522 TTTTCATGTTTTTCCCAAGTCACGAAGAGAGAAATGAGTTTATTTTGGCTCTGTAATGC 1581
Db 342 LeuArgPheLeu-----AspAspGlySerIleLeuTrpSerSerGlu--- 355
Qy 1582 AAAACAGGTTTCCGTCATTTATACAAAATTTACATCTATTTTAAAGGAAAGCAATATATAA 1641
Db 356 ArgThrGlyPheGlnHisLeuTyrArgIle-----AspSerLysGlyLys 370
Qy 1642 CGATCCAGTGGTGGCTGCTCTCCAAAGTATTCAAAGTATTCAGTGTCTATCAAGAGAGATA 1701
Db 371 AlaAla----- 372
Qy 1702 GCAATTACCACTGTAATGGGAAGTTCTTGGCGCGCATGGATCTAATATCAAGATTGAT 1761
Db 373 AlaLeuThrHisGlyAsnTrpSerVal-----AspGluLeuLeuAlaValAsp 388
Qy 1762 GAAGTCAGAAGGCTGGTATATTTTGAAGGCGACCAAGACTCCCTCTTTAGAGCATCACCTG 1821
Db 389 GluLysAlaGlyLeuAlaTyrPheArgAlaGlyIleGluSerAlaArgGluSerGlnIle 408

RESULT 4

S66261

X-Pro dipeptidyl-peptidase (EC 3.4.14.11) - Flavobacterium meningosepticum

C:Species: Flavobacterium meningosepticum

C:Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C:Accession: S66261
R:Kabashima, T.; Yoshida, T.; Ito, K.; Yoshimoto, T.
Arch. Biochem. Biophys. 320, 123-128, 1995
A:Title: Cloning, sequencing, and expression of the dipeptidyl peptidase IV gene from *Fluorobacterium*
A:Reference number: S66261; MUID:95314307; PMID:7793970
A:Accession: S66261
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-711 (KAB)
A:Cross-references: UNIPROT:Q47900; UNIPARC:UPI0000082115; EMBL:D42121; NID:G577283; PIR:1000000000
C:Superfamily: dipeptidyl-peptidase IV
C:Keywords: dipeptidylpeptide hydrolase

Alignment Scores:
Pred. No.: 7,38e-37 Length: 711
Score: 620.00 Matches: 195
Percent Similarity: 42.6% Conservative: 134
Best Local Similarity: 25.2% Mismatches: 284
Query Match: 11.2% Indels: 160
DB: 2 Gaps: 26

US-10-825-632-2 (1-3120) x S66261 (1-711)
QY 646 ATGGACAGTGGGATGCTTCTACGATTATCCACAGGA----- 687
Db 54 IIEGLUPRTHGLYILEALALYSYR-SERTYRYSRTHSRGlnLYSGlnLYSAsnILE 73
QY 688 ----AGTGGACATTT-----CTGTTTCAA 708
Db 74 VALASPGYSRPHGLNGLYTYR-PHESERASNAspGLUSerLYSILELeuLeuGln 93
QY 709 GCCGGTAGT-----GGATTTATCCAGTAAAGATGGA 741
Db 94 LYSER-SERGLN-SERILETYR-ARGHISER-PHELeuGLYLYSPHEGLUVALLYSAsp--- 112
QY 742 GGGCCCAAGGATTTACGCCAACACTTTAAGGCC---AATCTAGTGGNACTAGTTCT 798
Db 113 -----LeuLYS-SER-ARGHISER-PHEValSERLeuASN 123
QY 799 CCCAACATAGCGATGATCAAAATATGCCCCGTGATCCAGACTGGATGCTTTATA 858
Db 124 ALASNTripIleGlnLuproLYSPHESerPro---AspGLYSerLYSVALAlaPHEILE 142
QY 859 CATAGCACCATATTTGGATATCTAACATCGTAAACAGAGAAAGAGACTCACTTAT 918
Db 143 ALASpASNAsnLeuPheTYRGlnAspLeuAsnTHRGlyLYSILETHRGlnILETHR 162
QY 919 -----GTGCACAAATGAGTACCAACATGGAAGAGATGCCAGATCAGCTGGAGTCGCT 972
Db 163 ASPGLYLYSAsnGLUILEALEAN-----GlyLeuGly 174
QY 973 ACCTTTCTCCNAGAGAAATTTGATAGATATCTGGCTATTGGTGTGTCCAAAGCT 1032
Db 175 ASPTRPValTYRGlnGLUglnPHEGLYHISALAspTYRTYRGlnTRP----- 190
QY 1033 GAACAACATCCAGTGGTGTAAATTTCTAGAAATTTCTATATGAGAAATGATGAATCT 1092
Db 191 -----AsnLYSAlaGLYAspAlaLeuValPHEValARGPHEAspGLUARGLYSValPro 208
QY 1093 GAGTGGGAATTT---ATTGATGTATATCC-----CCTATGTTGAAACAGAGGGCA 1143
Db 209 GLUTILEAsnILEProIleTYRTYRGlnAsnLeuTYRProLYSLeuMetTHR----- 225
QY 1144 GATTCATTCCTTATCTAAACAGACAGTACAGCAATCTTAAGTCTTTTAAGATGCA 1203
Db 226 -----TYRLYSTYRProLYSAlaGLYglnGLUAsnSERAlaValTHRALATYRLeuTYR 243
QY 1204 GAAATATGATGATGCTGAGGAGGAGGATCATAGATGTCTATAGTAAGCAATATTCNA 1263
Db 244 GlnLeu-----SER-SERGLYLYS-----TATATTGCCAGAGCTGGAGT 1317
QY 1264 CCTTTTGAGATTTCTATTGAGGAGTTGAA-----TATATTGCCAGAGCTGGAGT 1317

Db 250 SerAlaGlnLeuAsnPheGlySerSerGlnLYSTYRILEProGlnLeuPheGlnThr 269
QY 1318 CCTGAGGAAATATGCTTGGTCCATCTACTAGATCGCTCCAGACTCGCTACAGATA 1377
Db 270 AsnAlaAsnAspGlnLeuValAlaThrAlaAsnARGHISGlnAsnLYSValAspLeu 289
QY 1378 GTGTTGATCTCAGCTCGAATTTATCCAGTAGAAGATGATGTTATGGAAGGACAGAA 1437
Db 290 LeuLYSValAsnTHRLYS----- 295
QY 1438 CTCATTGAGTCAGTCCGCTGATCTCTGACGCCACTAATATCTATGACAGACACACAC 1497
Db 296 -----ThrAlaAlaValSERLeuPheTHRLuTHRAspAsn 308
QY 1498 ATCTGGATAAATATCCATGACATCTTTTCATGTTTTTCCCAAGCTCACAGAGGAAT 1557
Db 309 AlaTRPileGLUthRAspAsnLeuTHRMETGLUPHE-----LeuAspAspAsn 324
QY 1558 GAGTTTATTTTGGCTCTGATGCAAAACAGGTTTCCGTCAATTATPACAAATATCATCT 1617
Db 325 SerPheLeuTPAlaSerGlu---ArgAspGLYHISARGHISLeuTYRTPTYRAspAla 343
QY 1618 ATTTTAAAGGAAGCAATATATAACGATCCAGTGGTGGCTGCTCTCCAAAGTGATTTC 1677
Db 344 AlaGLYLYSLeuLYSLYS----- 349
QY 1678 AAGTGTCTATCAAGAGGAGATAGCAATACCAGTGGTGAAATGGGAAGTTCTTGGCCGG 1737
Db 350 -----GlnValSERLYSGLYAspTRPGLUILEILEAsnTYR 361
QY 1738 CATGATCTATATATCAAGTTGATGAGTCAGAGGCTGATATTTTGAAGCACCAAA 1797
Db 362 TYRGLYTYRAsnProLYSTHRLYSGLU-----ValTYRILEGlnTHRTHRGlu 377
QY 1798 GACTCCCTTTAGACATCATCCTGTACTAGTCACT-----TACGTAAATCTGGAGAG 1851
Db 378 LYSGLYSERILEAsnLYS-----ValValSERLYSLeuAsnILEAsnTHRGlyLYS 394
QY 1852 GTGACAAAGCTGACTGACCGGTGCTACTCACTTCTGTGCTGATCAGTCACTGTGAC 1911
Db 395 THRGLNLeuLeuSERAsnAlaGLUGLYAsnAsnSERAlaAlaPHESerLYSTHRRPHEAsn 414
QY 1912 TCTTTATAGTAAGTATAGTAACCAAGAAATCCACACTGTGTGCTCTTACAGCTA 1971
Db 415 TYR-PHEILEAsnTHRSerSerThrAlaLYSValProTHRLYSTYRILELeuLYSAspAla 434
QY 1972 TCAAGTCTCGAA-----GATGACCAACTTGCACAAACAAAGGAA 2010
Db 435 AsnGLYLYSAspValLYSGlnLeuGlnAsnAspAspLeuLeuAsnLYSLeuLYS 454
QY 2011 TTTTGGGCCCACTTTTGGATTCCAGAGGCTCTCTCTGACTATCTCTCCAGAAAT 2070
Db 455 -----AspAsnPHEILEAlaLYSGluPHE 462
QY 2071 TTCTCTTTGAAAGTACTACTGATTTACATTGTTATGGATGCTCTACAGCCTCATGAT 2130
Db 463 ILETHRILEProAsnAlaAlaGLYAspGlnMetAsnAlaTRPMETILELYSPROLYSAsn 482
QY 2131 CTACAGCTGGAAGAAATATCTCTGCTGCTGTTTCATATATGTTGGTCTCTCAGTGAG 2190
Db 483 PHEAspProAlaLYSTYRProValPHEMetPHEGlnTYR-SERGLYProGLYSERGLN 502
QY 2191 TTGFGAATATCGTTTAAAGGAGTCAGTATTTCGCTTGAATACCTAGCTCTCTCTA 2250
Db 503 GlnValAlaAsnSERTRPAspGLYGLYAsnGLYILETRPHEAspMetLeuAlaGlnLYS 522
QY 2251 GGTATGTTGGTTGATGATAGACACAGGGGATCTCTGTCACCGAGGCTTAAATTTGAA 2310
Db 523 GLYTYRLeuValValCYsValAspGLYARGGLYTHRGlyPHEARGGLYTHRLYSTYR 542
QY 2311 GSGCGCTTTAAATATAAATGGGTCAATAGATAAATTTGACGATCAGTGGAGGACTCCNA 2370
Db 2370

Db 543 LysValThrTyrLysAsnLeuGlyLysTyrGluIleGluAspGlnIleThrAlaAlaLys 562
QY 2371 TATCTAGCTTCTCGATATGATTTCAATTGACTTAGATCGTGGGCATCCAGCGCTGCTCC 2430
Db 563 TrpLeuGlyAsnGln---SerTyrValAspLysSerArgIleGlyIlePheGlyTrpSer 581
QY 2431 TATCGAGCATACCTCTCCCTGATGGCATTAAATGCAGAGCTCAGATATCTTCAGCGTTGCT 2490
Db 582 TyrGlyGlyTyrMetAlaSerLeuAlaMetThrLysGlyAlaAspValPheLysMetGly 601
QY 2491 ATTGCTGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGAACTGTTAT 2550
Db 602 IleAlaValAlaProValThrAsnTrpArgPheTyrAspSerIleTyrThrGluArgPhe 621
QY 2551 ATGGGTCACCTCGACAGAAATGAACAGGCTATTACTTAGGATCTGTGGCATCCAGCA 2610
Db 622 LeuGlnThrProGlnGluAsnLysAspGlyTyrAspLeuAsnSerProThrThrTyrAla 641
QY 2611 GAAAGTTCCTCTGACCAATCGTTTACTGCTCTTACATGGTTCTGTCGATGAGAAT 2670
Db 642 LysLeuLeu-----LysGlyLysPheLeuLeuIleHisGlyThrAlaAspAsn 658
QY 2671 GTCCATTTTCACATACCAGTATATTACTGAGTGTTTTAGTGAGGCTGGAAAGCCATAT 2730
Db 659 ValHisPheGlnAsnSerMetGluPheSerGluAlaLeuIleGlnAsnLysLysGlnPhe 678
QY 2731 GATTTCACAGATCTATCTCAGGAGACACAGATAAGAGTTCCTGAATCGGAGAACAT 2790
Db 679 AspPheMetAlaTyrProAspLysAsnHisSerIleIleGlyGlyAsnThrArgProGln 698
QY 2791 TATGAACGTGATCTTTTGCACTACCTTCAAGAAAACCTT 2829
Db 699 LeuTyrGluLysMetThrAsnTyrIleLeuGluAsnLeu 711
RESULT 5
CDHU26
dipeptidyl-peptidase IV (EC 3.4.14.5) - human
N/Alternate names: cell surface glycoprotein CD26; thymocyte-activating molecule (THAM)
C/Species: Homo sapiens (man)
C/Date: 31-Dec-1993 #sequence revision 23-Aug-1996 #text change 09-Jul-2004
C/Accession: S24313; A42408; B61136; S59510; I56154; S59857; S15520
R/Misum, Y.; Hayashi, Y.; Arahawa, F.; Ikehara, Y.
Biochim. Biophys. Acta 1131, 333-336, 1992
A/Title: Molecular cloning and sequence analysis of human dipeptidyl peptidase IV, a serine protease
A/Reference number: S24313; MUID:92329551; PMID:1352704
A/Accession: S24313
A/Molecule type: mRNA
A/Residues: 1-6,'I',8-766 <MS>
A/Cross-references: UNIPROT:P27487; UNIPARC:UPI000016A580; EMBL:X60708; NID:g35335; PIDN:RjDatmoul, D.; Lacasa, M.; Baricault, L.; Marguet, D.; Sapin, C.; Trotot, P.; Barbat, A. J. Biol. Chem. 267, 4824-4833, 1992
A/Title: Dipeptidyl peptidase IV (CD 26) gene expression in enterocyte-like colon cancer cell lines during cell differentiation.
IV mRNA levels during cell differentiation.
A/Reference number: A42408; MUID:92165847; PMID:1347043
A/Accession: B42408
A/Molecule type: mRNA
A/Residues: 1-5,'R',7-436,'S',438-556,'I',558-662,'E',664-766 <DAR1>
A/Cross-references: UNIPARC:UPI0000052ACB; GB:M80536; NID:g181569; PIDN:AAA52308.1; PID:J. Biol. Chem. 267, 4824-4833, 1992
A/Experimental source: intestine
A/Note: this sequence corresponds with the author's translation
A/Accession: A42408
A/Molecule type: mRNA
A/Residues: 1-5,'R',7-436,'S',438-556,'I',558-662,'E',664-711,'G',713-766 <DAR2>
A/Cross-references: UNIPARC:UPI0000172A2B; GB:M80536; NID:g181569
A/Note: sequence extracted from NCBI backbone (NCBIIN:83986, NCBIIP:83988); this sequence R/Gorvel, J.P.; Ferrero, A.; Chambraud, L.; Rigal, A.; Bonicel, J.; Maroux, S. Gastroenterology 101, 618-625, 1991
A/Title: Expression of sucrase-isomaltase and dipeptidylpeptidase IV in human small intestine
A/Reference number: A61136; MUID:91317403; PMID:1677636
A/Accession: B61136
A/Molecule type: protein
A/Residues: 1-15,'X',17-22 <GOR>
A/Cross-references: UNIPARC:UPI0000172A2C

RjBoehm, S.K.; Gum Jr., J.R.; Erickson, R.H.; Hicks, J.W.; Kim, Y.S. Biochem. J. 311, 835-843, 1995
A/Title: Human dipeptidyl peptidase IV gene promoter: tissue-specific regulation from a 1.5 kb promoter region
A/Reference number: S59510; MUID:96067939; PMID:7487939
A/Accession: S59510
A/Molecule type: DNA
A/Residues: 1-31 <BOB>
A/Cross-references: UNIPARC:UPI000016B4A6; GB:S79876; NID:g1195574; PIDN:AA835614.1; PID:RjTanaka, T.; Camerini, D.; Seed, B.; Torimoto, Y.; Dang, N.H.; Kameoka, J.; Dahlberg, J. Immunol. 149, 481-486, 1992
A/Title: Cloning and functional expression of the T cell activation antigen CD26.
A/Reference number: I56154; MUID:92325476; PMID:1352530
A/Accession: I56154
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-436,'S',438-766 <TAN>
A/Cross-references: UNIPARC:UPI000004F7BF; GB:M74777; NID:g180082; PIDN:AAA51943.1; PID:RjAbbott, C.A.; Baker, E.; Sutherland, G.R.; McCaughan, G.W. Immunogenetics 40, 331-338, 1994
A/Title: Genomic organization, exact localization, and tissue expression of the human CD26.
A/Reference number: S59857; MUID:95012454; PMID:7927537
A/Accession: S59857
A/Molecule type: DNA
A/Residues: 1-436,'S',438-766 <ABB>
A/Cross-references: UNIPARC:UPI000004F7BF; EMBL:U13734
C/Genetics:
A/Gene: GDB:DPP4
A/Cross-references: GDB:125239; OMIM:102720
A/Map position: 2q24.3-q24.3
A/Introns: 2/3; 32/1; 65/1; 95/3; 122/3; 140/2; 164/3; 205/1; 258/3; 296/2; 341/3; 356/3
C/Superfamily: dipeptidyl-peptidase IV
C/Keywords: dipeptidylpeptidase hydrolase; glycoprotein; homodimer; proteinase; transmembrane protein
F1-6/Domain: intracellular #status predicted <INT>
F7-28/Domain: transmembrane #status predicted <TMN>
F29-766/Domain: extracellular #status predicted <EXT>
F85,92,150,219,229,281,321,520,685/Binding site: carbohydrate (Asn) (covalent) #status predicted
F630,708,740/Active site: Ser, Asp, His #status predicted
Alignment Scores:
Pred. No.: 2,97e-30 Length: 766
Score: 529.00 Matches: 186
Percent Similarity: 42.0% Conservative: 109
Best Local Similarity: 26.5% Mismatches: 280
Query Match: 9.5% Indels: 128
DB: 1 Gaps: 32
US-10-825-632-2 (1-3120) x CDHU26 (1-766)
QY 829 CCCTCATCCAGCTGGATT-----GCTTTTATACAT 861
Db 149 ProAsnAsnThrGlnTrpValThrTrpSerProValGlyHisLysLeuAlaTyrValTrp 168
QY 862 AGCAACATATTTGGATATCTACATCGTAACACAGAGAAGAAAGGAGCTCACTTATGTG 921
Db 169 AsnAsnAspIleTyrValLysIleGluProAsnLeuProSerTyrArgIleThrThr 188
QY 922 CACAATGAGTAGCAACATCGAAGAGAGATGCGCAGCTGAGTGCCTACCTTTGTT 981
Db 189 Gly-----LysGluAspIleIleTyrAsnGlyIleThrAspTrpVal 202
QY 982 CTCCAAGAAGAA---TTTGATAGATATCTGCGTATTGGTGTGTCCTCAAAAGCTGAAACA 1038
Db 203 TyrGluGluGluValPheSerAlaTyrSerAlaLeuTyrTrpSerProAsnGlyThrPhe 222
QY 1039 ACTCCAGTGGTGAATAATCTTGAATTCATATGAGAAATAATGATGAATCTGAGGTG 1098
Db 223 -----LeuAlaTyrAlaGlnPheAsnAspThrGluVal 233
QY 1099 GAAATATTTCAT-----GTTACATCCCCTATGTTGGAAACAGAGGCGCAGATCA 1149
Db 234 ProLeuIleGluTyrSerPheTyrSerAspGluSerLeuGlnTyrProLysThrValArg 253
QY 1150 TTCGGTATCCTAAACAGGTACAGCAAAATCCTAAAGTCACCTTTTAAAGATGTCAGAAATA 1209

Db 254 ValProTyrProLysAlaGlyAlaValAenProThrValLysPhePheVal-----Val 271
 QY 1210 ATGATTCATCTCAAGAGGATCATAGATGTATAGATAGAGAACTAAATCAACCTTTT 1269
 Db 272 AenThrAspSerLeuSerValThrAenAlaThrSerIleGlnIleThrAlaProAla 291
 QY 1270 GAGATTCATTTGAAGGAGTGAATATATATGTCAGAGCTGGATGGACTCTCTGAGGAAAA 1329
 Db 292 SerMetLeuLeu-----GlyAspHisTyrLeuCysAspValThrTrpAla----- 306
 QY 1330 TATGCTGGTCCCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCA 1389
 Db 307 -----ThrGlnGluArgIleSerLeuGlnTrpLeu--- 316
 QY 1390 CCGTGAATATTATCCAGTAGAAGATGATGTTATGGAAGACAGAGACTCAATGAGTCA 1449
 Db 317 -----ArgArgIleGlnAen 321
 QY 1450 GTGCTGATTCGTGAGCCCACTAATATCTATGAAGAACACAGACATCTGGATAAT 1509
 Db 322 Tyr-----SerValMetAspIleCysAspTyrAspLysSerSerGlyArgTrp---Aen 338
 QY 1510 ATCCATGATCTTTTCATGTTTTCCCAAGTTCACAGAGAGAAATTCAGTTATTTT 1569
 Db 339 Cys-----LeuValAlaArgGlnHisIleGluMetSerThrThrGlyTrp 353
 QY 1570 GCCTCTGAATGCMAAACAGGTTTCCGTCAT-----TTATACAA 1608
 Db 354 ValGlyArgPheArgProSerGluProHisPheThrLeuAspGlyAenSerPheTyrLys 373
 QY 1609 ATTATCATCTATTATTAAGAGAAAGCAATATAAAGCATCCAGTGGTGGCTGCTGCTCA 1668
 Db 374 IleIleSer-----AsnGluGluGlyTyrArgHisIle----- 384
 QY 1669 AGTGATTTCAAGTGCTCTCAAGAGGAGATAGCAATATACAGTGGTGAATGGAGTT 1728
 Db 385 CysTyrPheGlnIleAspLysLysAspCysThrPheIleThrLysGlyThrTrpGluVal 404
 QY 1729 CTGGCGGCATGGATTAATATCAAGTTGATGAAGTTCAGAGAGCTGGTATATTTGAA 1788
 Db 405 IleGly-----IleGluAlaLeuThrSerAspTyrLeuTyrTyrIleSer 419
 QY 1789 GGCACC-----AAAGACTCCCTTTAGAGCATCACCTGACGTAGTACGTAAATCTC 1845
 Db 420 AsnGluTyrLysGlyMetProGlyArgAenLeuTyrLysIleGlnLeuIleAspTyr 439
 QY 1846 GAGAGGTGACAGGCTGACTACCGTCTACTCACATTTCTTCTGCTGACAGT---CAG 1902
 Db 440 ThrLysValThrCysLeu-----SerCysGluLeuAenProGlu 452
 QY 1903 CACTGTGACTCTTTTATAGTATAGTATAGTACCAAGAGATCCACAC-----TGT 1953
 Db 453 ArgCysGlnTyrTyrSerValSerPheSerLysGluAlaLysTyrTyrGlnLeuArgCys 472
 QY 1954 -----GTGCTCCCTTTCAAGCTATCAAGTCTGAGATGAGCACTGCAAA 2001
 Db 473 SerGlyProGlyLeuProLeuTyrThrLeuHisSerSerValAenAspLysGlyLeuArg 492
 QY 2002 ACAAGGAATTTTGGGCCCACTTTTGGATTCAGAGGCTCTCTTCTGACTATCTCTCT 2061
 Db 493 ValLeuGluAsp-----AenSerAlaLeuAspLysMet-----LeuGlnAenValGlnMet 509
 QY 2062 CCAGAA-----ATTTCTCTTTTGAAGTACTGAGTATTTACATTTGATGGATG 2112
 Db 510 ProSerLysLysLeuAspPheIleLeuAenGluThrLysPhe---TrpTyrGlnMet 528
 QY 2113 CTCTACAGCTCATGATCTACAGCTGGAAGAAATATCTTACTGTGCTGTCTATATAT 2172
 Db 529 IleLeuProProHis-----PheAspLysSerLysLysTyrProLeuLeuAspValTyr 547
 QY 2173 GGTGCTCTCTCAGGTGACGTTGGTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTG 2232

Db 548 AlaGlyProCysSerGln-----LysAlaAspThrValPheArgLeu 561
 QY 2233 AAT-----ACCCTAGCTCTCTAGGTATATGTTGTAGTG---ATAGACACAGG 2280
 Db 562 AenTrpAlaThrTyrLeuAlaSerThrGluAenIleIleValAlaSerPheAspGlyArg 581
 QY 2281 GGATCTCTCAACGAGGCTTAAATTTGAAGCGCTTTAAATATAATAATGGGTCAATA 2340
 Db 582 GlySerGlyTyrGlnGlyAspLysIleMetHisAlaIleAenArgArgLeuGlyThrPhe 601
 QY 2341 GAAATTCACCATCAGGTGGAAGGACTCCCAATATCTAGTCTTCGATATGATTTCAATGAC 2400
 Db 602 GluValGluAspGlnIleGluAlaAaGlnPhe---SerLysMetGlyPheValAsp 620
 QY 2401 TTAGATCGTGGGCATCCAGGCTGCTCTATGAGGATACCTCTCCCTGATGGCATTA 2460
 Db 621 AsnLysArgIleAlaIleTrpGlyTyrSerTyrGlyTyrValThrSerMetValLeu 640
 QY 2461 ATGCAGAGTCCAGTATCTTCAGGTTGCTATTGCTGGGGCCCACTCTGTGGATC 2520
 Db 641 GlySerGlySerGlyValPheLysCysGlyIleAlaValAlaProValSerArgTrpGlu 660
 QY 2521 TTCATATGATACAGATACAGGACCTTATATGGT-----CACCTGACCAAGATGAA 2574
 Db 661 TyrTyrAspSerValTyrThrGluArgTyrMetGlyLeuProThrProGluAspAenLeu 680
 QY 2575 CAGGCTATTACTAGATCTGGCCATCCAGAGGAGAAAGTTCCCTCTGAACCAAT 2634
 Db 681 AspHisTyrArgAenSerThrValMetSerArgAlaGluAenPhe-----LysGlnVal 698
 QY 2635 CGTTTACTGCTCTTACATGTTTCTCGATGAGATGTCATTTCCATTTGCACATACAGTATA 2694
 Db 699 GluTyrLeuLeuIleHisGlyThrAlaAspAenValHisPheGlnGlnSerAlaGln 718
 QY 2695 TTACTGAGTTTTATGAGGCTGGAAGCCATATGATTACAGATCTATCTCAGGAG 734
 Db 719 IleSerLysAlaLeuValAspValGlyValAspPheGlnAlaMetTrpTyrThrAspGlu 738
 QY 2755 AGACACAGCATTAAGAGTTCCTGAATCGGGAGAACATTATGAATGCTTTTGGCTACTAC 2814
 Db 739 AspHisGlyIleAlaSerSerThrAlaHisGlnHisIleTyrThrHisMetSerHisPhe 758
 QY 2815 CTTCAAGAA 2823
 Db 759 IleLysGln 761
 RESULT 6
 T41703 dipeptidyl aminopeptidase - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
 R:Murphy, L.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
 Submitted to the EMBL Data Library, August 1998
 A:Reference number: Z22011
 A:Accession: T41703
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-793 <MUR>
 A:Cross-references: UNIPROT:O14073; UNIPARC:UPI000013A955; EMBL:AL031180; PIDN:CAA20138
 A:Experimental source: strain 972h-; cosmid c2511 -chimeric
 C:Genetics:
 C:Gene: SPAC2E11.08
 A:Map position: 1
 C:Superfamily: dipeptidyl-peptidase IV
 Alignment Scores:
 Prod. No.: 1.46e-29 Length: 793
 Score: 519.50 Matches: 187
 Percent Similarity: 42.2% Conservatives: 113
 Best Local Similarity: 26.3% Mismatches: 256
 Query Match: 9.4% Indels: 155
 DB: 2 Gaps: 32

US-10-825-632-2 (1-3120) x T41703 (1-793)

QY	847	ATTGCTTTTATACATAGCAACCATATTTGGATA-----TCTAACATCGTAACC	894
Db	183	LeuSerPheValTyraAsnAseLeuTyraValaLysAsnAseGlyAsnVal----	200
QY	895	AGAGAAGAAGGAGACTCCTTAT-----GTGCACATGACTAGCCAAAC	939
Db	201	-----GlnArgLeuThrTyraAspGlyThrAlaValPheAsn-----	213
QY	940	ATGGAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAGAAGAA--TTT	996
Db	214	-----GlyLeuThrAspTrpIleTyrgluGluGluValLeu	225
QY	997	GATAGATATTTCTGGCTATTGGTGTTCCAAAAGCTGAACAACATCCCGTGGTGA	1056
Db	226	SerSerProSerThrIleTrpSerProAspSerAsp-----	238
QY	1057	ATTCTTAGAATCTTATGAAGAATATGATGATCTGAGGTGGAAATATTTCATGTT---	1113
Db	239	-----LysIleAlaPheLeuLysLeuAsnGluSerGluIleProThrTyHisTyPro	256
QY	1114	-----ACATCCCCTATGTTGGAAACAAAGGAGGCGCAGATTCA	1149
Db	257	LeuTyThrAlaGluLeuAspProSerLeuProGluPheAspTyraSnLysAspMetAla	276
QY	1150	TTCGGTATCTCTAAAAACAGTACAGCAAATCTCTAAAGTCACCTTTTAAGATGTCAGAAATA	1209
Db	277	IleLysTyProLysProGlyAsnProAsnProSerValSerLeuPheValAlaAspLeu	296
QY	1210	ATGATTGATGCTGAAGGAAGGATCATAGATGTCTATAGATAAGAACTAATTCAACCTTTT	1269
Db	297	AsnSerAsnAlaSerSerAsnPheSerLeuTrpHisAsnGluProLeuAlaGluPro---	315
QY	1270	GAGATCTATTGGAAGGATTTGAATATATATGTCAGAGCTGGATGACTCTCGAGGGA	1329
Db	316	-----ValValGlnAsnValLeuTrpValAsnThrSer-----	326
QY	1330	TATGCTTGGTCCATCTCTACTA-----GATCGCTCCCAGACTCGCTACAGATAGTG	1380
Db	327	-----SerValLeuValGlnPheThrAsnArgAsnSerThrCysIleThrAlaArg	343
QY	1381	TTGATCTCACCTCAATTA-----TTTATCCAGTAGAAGATGATGTTATGGAAGG---	1431
Db	344	LeuLeuAspThrGluLeuLysSerIleHisThrValLysThrGluCysLeuGluGluGly	363
QY	1432	-----CAGAGACTCATAGTCAGTCAGTCCCTGATCTGTGAGCCCACTAATTATC	1479
Db	364	TriTyr-GluValGlnGlnSerAlaLysMetPhePro-----LeuAsnAsnSerLeuVal	381
QY	1480	TATGAAGAACAACACACATCTGGATAAATATCCATGACATCTTTCATGTTTTCCCCAA	1539
Db	382	TriGluAsnTrpSerAsp-----GlyTyrPheAspIleLeuAlaLeuAspAspTy	398
QY	1540	AGTCACGAAGAGAAATTCAGTTTATTTTGGCTCTGAAATGCAAAACAGGTTTCCGTCAT	1599
Db	399	AsnHisLeuAlaPheIleProPhe-----	406
QY	1600	TTATACAAAATTATCATCTATTTTAAAGGAAGCAAAATATAAACGATCCAGTGGGCTG	1659
Db	407	-----AsnGlySerSer	410
QY	1660	CCTGCTCAAGTGATTTCAAGTGTCTATCAAGAAGGAGATAGCAATTACCAGGTGGAA	1719
Db	411	Pro-----IleTyrLeuThrSerGlyAla	418
QY	1720	TGGGAAGTCTCTGGCGGCATGGATCTAATATCCAAAGTTGATGAAGTCAGAAGCTGGTA	1779
Db	419	TrpAspValThr-----aspGlyProIleHisIleAspGlyAspPheGlyAsnVal	435
QY	1780	TATTTTGAAGGCCAACCAAGACTCCCTTTTAGAGCATCACCTGTAGTACGTAGTACGTGA	1839
Db	771	SerGluTrpIleGlyAspAlaLeuGlyArgile	781
QY	2818	CAAGAAAACCTTGGATCACGATTATGCTGCTCTA	2850
Db	771	SerGluTrpIleGlyAspAlaLeuGlyArgile	781

RESULT 7
 S23752 dipeptidyl-peptidase IV (EC 3.4.14.5) alpha chain - mouse
 N:Alternate names: CD26 alpha subunit; THEM alpha subunit
 C:Species: Mus musculus (house mouse)
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C:Accession: S23752; A46465; A56030
 R:Marguet, D.; Bernard, A.M.; Vivier, I.; Darmoul, D.; Naquet, P.; Pierres, M.
 J. Biol. Chem. 267, 2200-2208, 1992
 A:Title: cDNA cloning for mouse thymocyte-activating molecule. A multifunctional ecto-di-
 A:Reference number: S23752; MUID:92129288; PMID:1370813
 A:Accession: S23752
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-760 <MAR>
 A:Cross-references: UNIPARC:UPI0000172A2D; EMBL:X59384
 R:Vivier, I.; Marguet, D.; Naquet, P.; Bonicel, J.; Black, D.; Li, C.X.; Bernard, A.M.;
 J. Immunol. 147, 447-454, 1991
 A:Title: Evidence that thymocyte-activating molecule is mouse CD26 (dipeptidyl peptidase
 A:Reference number: A46465; MUID:91302787; PMID:1712807
 A:Accession: A46465
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-20 <VIV>
 A:Cross-references: UNIPARC:UPI0000172A2E
 A:Experimental source: M14 T thymoma cells, Swiss nu/nu
 A:Note: sequence extracted from NCBI backbone (NCBIP:42236)
 R:Bernard, A.M.; Mattei, M.G.; Pierres, M.; Marguet, D.
 Biochemistry 33, 15204-15214, 1994
 A:Title: Structure of the mouse dipeptidyl peptidase IV (CD26) gene.
 A:Reference number: A56030; MUID:95092780; PMID:7999781
 A:Accession: A56030
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 746-760 <BBR>
 A:Cross-references: UNIPARC:UPI0000172A2F; GB:U12620
 C:Genetics:
 C:Superfamily: dipeptidyl-peptidase IV
 C:Keywords: dipeptidyl-peptide hydrolase; glycoprotein; transmembrane protein
 F:213,223,315,514,679/Binding site: carbohydrate (Asn) (covalent) #status Predicted
 F:624,702,734/Active site: Ser, Asp, His #status Predicted

Alignment Scores:
 Pred. No.: 1,71e-29 Length: 760
 Score: 518.50 Matches: 203
 Percent Similarity: 38.1% Conservative: 107
 Best Local Similarity: 25.0% Mismatches: 306
 Query Match: 9.3% Indels: 197
 DB: 1 Gaps: 37

US-10-825-632-2 (1-3120) x S23752 (1-760)
 QY 604 ATGTATCTCGAAGNAGAACTATTAAAGAGAAAGAAACGATTGGAAACAGTCGGAATT 663
 Db 67 LeuTyrLysGlnGluAsnAsnIleLeuLeuLeuAsnAlaGluHisGlyAsnSerSerIle 86
 QY 664 -----GCTTCTTACGATTATCACCAGGAAGTGGAAACATTTCG 702
 Db 87 PheLeuGluAsnSerThrPheGluSerPheGlyTyrHisSerValSerProAspArgLeu 106
 QY 703 TTT -----CAAGCCGGTAGT 717
 Db 107 PheValLeuLeuGluTyrAsnTyrValLysGlnTrpArgHisSerTyrThrAlaSerTyr 126
 QY 718 GGAATTTATCAGCTAAAGATGAGGGCCCAAGGATTTACCAACAACCTTTAAGGCC 777
 Db 127 AsnIleTyrAspValAsnLys -----ArgGln 135
 QY 778 AATCTAGTGGAAACTAGTTCTCCCAACATACCGATGGATCCAAATATTATGCCCGCTGAT 837
 Db 136 LeuIleThrGluGluLysIleProAsn-----Asn 145

QY 838 CCAGACTCGATT-----GCTTTATATACATAGCAACGAT 870
 Db 146 ThrGlnTrpIleThrTrpSerProGluGlyHisLysLeuAlaTyrValTrpLysAsnAsp 165
 QY 871 ATTTGGATATCTAAACATCGTAACACAGAGAAGAAGAGAGACTCATTATGTGCACATGAG 930
 Db 166 IleTyrValLysValGluProHisLeuProSerHisArgGlnThr-----180
 QY 931 CTAGCCAACTATGGGAAGAGATGCAGATCAGCTCGGAGTCGCTACCTTTTCTCCAGAA 990
 Db 181 ---SerThrGlyGluGluAsnValIleTyrAsnGlyIleThrAspTrpValTyrGluGlu 199
 QY 991 GAA-----TTTGATAGATATTTCTGGCTATTTGGTGTGTCCAAA-----1029
 Db 200 GluValPheGlyAlaTyrSerAlaLeuTrpTrpSerProAsnAsnThrPheLeuAlaTyr 219
 QY 1030 GCTGAAACAACTCCCGAGTGGTAAATAATCTTAGAATCTATATAGAGAAATGATGAA 1089
 Db 220 AlaGlnPheAsnAspThrGlyValProIleGluTyrSerPhe---TyrSerAspGlu 238
 QY 1090 TCTGAGGTGGAAATTTATTCATGTTTATCATCCCTCTGTTGGAACAAGAGGGCAGATTCA 1149
 Db 239 Ser-----LeuGlnTyrProLysThrValTrp 247
 QY 1150 TTCGGTTATCTTAAACAGGTACAGCAATCTTAAAGTCACTTTTAAAGATGTGCAGAAATA 1209
 Db 248 IleProTyrProLysAlaGlyAlaValAsnProThrValLysPhePheIle-----Val 265
 QY 1210 ATGATTGATGCTGAAGGAGGATCATAGATGTCATAGTAAGGAACATAATCAACCTTTT 1269
 Db 266 AsnIleAspSerLeuSerSerSerSerSerAlaProIleGlnIleProAlaProAla 285
 QY 1270 GAGATTCTATTTGAAGGAGTGAATATATTCGCGAGCTGGATGAGTCTCTCGAGGAAA 1329
 Db 286 SerValAla---ArgGlyAspHisTyrLeuCysAspValValTrpAlaThrGluGluArg 304
 QY 1330 TATGCT-----TGGTCCATCTTACTAGATCGCTCCAG 1362
 Db 305 IleSerLeuGlnTrpLeuArgIleGlnAsnTyrSerValMetAlaIleCysAspTyr 324
 QY 1363 ACTCGCTACAGATAGTGTGATCTCAGCTGAATTTATTTATCCAGTAGAAGATGATGTT 1422
 Db 325 AspLysIleAsnLeuThrTrpAsnCysProSer-----335
 QY 1423 ATGMAAGGACAGAGACTCATTGAGTCAGTCGCTGATCTGTGACGCCACTAATTATCTAT 1482
 Db 336 ---GluGlnGlnHisVal-----340
 QY 1483 GAAGAACAACACACATCTGGATAAATATCCATGACATCTTTTCATGTTTTCCTCCAAAGT 1542
 Db 341 GluMetSerThrThrGlyTrpValGly-----AspPheArgProAlaGluProHisPhe 358
 QY 1543 CACGAGAGAAATTGAGTTTATTTTGGCTCTGAATGCACAAACAGGTTTCCGTCATTTA 1602
 Db 359 ThrSerAspGlySerSerPheTyrLysIleIleSerAspLysAspGlyTyrLysHisIle 378
 QY 1603 TACAAATTCATCTATTTTAAAGGAAAGCAATATATAACGATCCAGTGGTGGCTGCCT 1662
 Db 379 CysHisPhe-----381
 QY 1663 GCTCCAAAGTGATTTCAG-----TGTCTATCAAGAGAGATAGCAATACCAAGTGGTGA 1719
 Db 382 ---ProLysAspLysLysAspCysThrPhe-----IleThrLysGlyAla 395
 QY 1720 TGGGAAGTTCTTTCGCCCGCATGGATCTAATATCAAGTTGATGAGTCAAGAGGCTGTA 1779
 Db 396 TrpGluValIle-----SerIleGluAlaLeuThrSerAspTyrLeuTyr 410
 QY 1780 TATTTTGAAGGCCACC-----AAAGACTCCCTTTAGAGCATCAGCTGTGACGTAGTAC 1836
 Db 411 TyrIleSerAsnGlnTyrLysGluMetProGlyGlyArgAsnLeuTyrLysIle-----428

F:83,90,148,217,227,319,521,686/Binding site: carbohydrate (Asn) (covalent) #status pred
 F:631/Active site: Ser #status experimental
 F:709,741/Active site: Asp, His #status predicted

Alignment Scores:
 Pred. No.: 6,57e-29 Length: 792
 Score: 510.50 Matches: 182
 Percent Similarity: 40.6% Conservative: 105
 Best Local Similarity: 25.7% Mismatches: 287
 Query Match: 9.2% Indels: 133
 Ds: 1 Gaps: 30

US-10-825-632-2 (1-3120) x A39914 (1-792)

Qy	829	CCCGTATCCAGACTGGATGCTTTATACATAGCAACGAT-----ATTGG	876
Ds	147	ProAsnThrGlnTrpSerGlnGluGlyHisLeuAlaTrpValTrp	166
Qy	877	ATATCAATCATCGTAACAGAGAAAGAGACTCATTATGTGCACATAGCTAGCC	936
Ds	167	LysAsnAspIleTyrValLysIleGluProHisLeuPro---SerHisArgThrThrSer	185
Qy	937	AACATGGAAGAGATGCCAGATCAGCTGGAGTCGCTACCTTGTCTCCAGAGAA---	993
Ds	186	ThrGlyLysGluAsnValIlePheAsnGlyIleAsnAspTrpValTyrGluGluGluLe	205
Qy	994	TTTGATAGATATTCCTGGCTATTTGGTGTGTCCAAAAGCTGAACAACTCCCACTGGT	1053
Ds	206	PheGlyAlaTyrSerAlaLeuTrpTrp-----SerProAsnGlyThr	219
Qy	1054	AAATATCTTAGAATTTCTATATGAGAAATAATGATCAATCTGAGTGGAATATTATCATGTT	1113
Ds	220	PheLeuAlaTyrAlaGlnPheAsnAspThrGlyValProLeuIleGluTyrSerPheTyr	239
Qy	1114	ACATCCCTATGTTGGAAACAAGAGGCGAGATTCCTCGTTATCCTTAAACAGGTACA	1173
Ds	240	SerAspGluSerLeuGlnTyrProLysThrValTrpIleProTyrProLysAlaGlyAla	259
Qy	1174	GCARATCTTAAGTCACCTTTTAAGATCTCAGAAATATGATTCATCTCAAGAGGATC	1233
Ds	260	ValAsnProThrValLysPhePheIle-----ValAsnThrAspSerLeuSerThr	277
Qy	1234	ATAGATGTCATAGATAGGAACATAATTCACCTTTTGAGATTCTATTGAGGAGTTGAA	1293
Ds	278	ThrThrThrIleProMetGlnIleThrAlaProAlaSerValThr---ThrGlyAspHis	296
Qy	1294	TATATGCCAGCTGGATGAGCTCCCTGAGGAAATATGCTTGGTCCATCTCTACTAGAT	1353
Ds	297	TyrLeuCysAspValAlaTrpValSerGluAspArgIleSer-----	310
Qy	1354	CGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTCGAATATTATCCAGTAGAA	1413
Ds	310	-----	310
Qy	1414	GATGATGTTATGGAAGCGAGAGACTCATTGAGTCAGTCGCTGATCTGTGAGCGCACTA	1473
Ds	311	-----LeuGlnTrpLeuArgIleGlnAsnTyr-----SerValMetAlaIle	325
Qy	1474	ATTATCTATGAAGAAACAACAGACTCTGGATAAATATCCATGACATCTTTTCATGTT---	1530
Ds	326	CysAspTyrAspLysThrAsnLeuValTrpAsnCysProThrThrGlnGluHisIleGlu	345
Qy	1531	-----TTTCCCAAGTCCAGGAAG	1551
Ds	346	ThrSerAlaThrGlyTrpValGlyArgPheArgProAlaGluProHisPheThrSerAsp	365
Qy	1552	GAAATGATGTTATTTTTCCTCGATGCAAAACAGTTTCCGTCTATTATACAAATTT	1611
Ds	366	GlySerSerPheTyrLysIleValSerAspLysAspGlyTyrLysHis-----Ile	382
Qy	1612	ACATCTATTTTAAAGGAAGCAATATAACAGTCCAGTGTGGCTGCTGCCAAGT	1671
Ds	383	CysGlnPheGlnLysAspArgLys-----	390

Qy	1672	GATTTCAAGTGTCTCTATCAAGAGAGAGATAGCAATTTACCAAGTGTGTAATGGGAGTCTT	1731
Ds	391	-----ProGluGlnAspCysThrPheIleThrLysGlyAlaTrpGluValIle	406
Qy	1732	GGCCGGCATGATCTTAATATCAAGTTGATGAGTGCAGAAAGCTGGTATATTTTGAAGGC	1791
Ds	407	-----SerIleGluAlaLeuThrSerAspTyrLeuTyrTyrIleSerAsn	421
Qy	1792	ACC---AAAGACTCCCTTTAGAGCATCACCTGTACGTAGTACGTACGTAAATCTCTGGA	1848
Ds	422	GluTyrLysGluMetProGlyGlyArgAsnLeuTyrLysIle---	435
Qy	1849	GAGGTGACAAAGCTGACTGACCTGGCTACTACATCTCTGCTGCATCACT-----	1899
Ds	436	-----GlnLeuThrAsp-----HisThrAsnLysLysCysLeuSerCysAspLeu	450
Qy	1900	-----CAGCACTGTCTCTT-----ATAAGTAAGTATATACCAACAGAGAAT	1944
Ds	451	AsnProGluArgCysGlnTyrTyrSerValSerLeuSerLysGluAlaLysTyrTyrGln	470
Qy	1945	CCACACTGT-----GTGCTCTTTTACAAGCTATCAAGTCTCTGAAGATGACCCA	1992
Ds	471	LeuGlyCysArgGlyProGlyLeuProLeuTyrThrLeuHisArgSerThrAspGlnLys	490
Qy	1993	ACTTGCAAAACAAGAAATTTGGCCACCATTGGATTCAGCAGCTCTCTCTCTGAC	2052
Ds	491	GluLeuArgValLeuGluAsp---AsnSerAlaLeuAspLysMet-----LeuGlnAsp	507
Qy	2053	TATACTCTCCAGAA-----ATTTCTCTTTTGAAGTACTACTGATTTACATTTG	2103
Ds	508	ValGlnMetProSerLysLeuAspPheIleValLeuAsnGlnThrArgPhe---Trp	526
Qy	2104	TATGGATCTCTCAAGCCTCATGATCTACAGCCTGGAAAGAAATATCTCTACTGCTG	2163
Ds	527	TyrGlnMetIleLeuProProHis---PheAspLysSerLysLysTyrProLeuLeuIle	545
Qy	2164	TTCATATATGGTGTCTCTCAGTGCAGTGTGTGTAATAATCGTTTAAAGAGTCAAGTAT	2223
Ds	546	AspValTyrAlaGlyProCysSerGln-----LysAlaAspAlaAla	559
Qy	2224	TTCCGCTTCAAT-----ACCTAGCTCTCTAGTATTATGTTGGTTGATGTG---ATA	2271
Ds	560	PheArgPheAsnTrpAlaThrTyrLeuAlaSerThrGluAsnIleIleValAlaSerPhe	579
Qy	2272	GACAAAGGGGATCTGTCCAGCGGCTTAAATTTGAAGCGCTTTAAATATAAAATG	2331
Ds	580	AspGlyArgGlySerGlyTyrGlnGlyAspLysIleMetHisAlaIleAsnLysArgLeu	599
Qy	2332	GGTCAATAGAAATTCAGATCAGCTGAGGAGGACTC---CAATATCTAGCTTCTCGATAT	2388
Ds	600	GlyThrLeuGluValGluAspGlnIleGluAlaAlaArgGlnPheLeu-----LysMet	617
Qy	2389	GATTTCAATGATCTAGTCTGTGGCATCCAGCTGCTCTATGGAGGATACCTCTCC	2448
Ds	618	GlyPheValAspSerLysGlnValAlaIleTrpGlyTrpSerTyrGlyTyrValThr	637
Qy	2449	CTGATGGCATTAATGAGAGGTACATATCTTCAGGTTGCTATTTGCTGGGCGCCAGTC	2508
Ds	638	SerMetValLeuGlySerGlySerGlyValPheLysCysGlyIleAlaValAlaProVal	657
Qy	2509	ACTCTGTGATCTCTTATCATACAGGATACACGGAACCTTATATGGGT-----CACCT	2562
Ds	658	SerArgTrpGluTyrTyrAspSerValTyrThrGluArgTyrMetGlyLeuProThrPro	677
Qy	2563	GACCAAGTAACAGAGGCTATTACTTAGCATCTGGCCATGCAAGCAAGAAAGTTCCCC	2622
Ds	678	GluAspAsnLeuAspHisTyrArgAsnSerThrValMetSerArgAlaGluAsnPhe---	696
Qy	2623	TCTGAACCAAAATCGTTTACTGCTCTTACATGTTTCTCTGGATGAGAATGTCCATTTTGA	2682
Ds	697	---LysGlnValGluTyrLeuLeuIleHisGlyThrAlaAspAsnValHisPheGln	715

```
QY 2683 CATACAGATATTAATGAGTTTCTAGTGGGCTGGAAGCCATATGATTTACAGATC 2742
Db      :::::  ::::  |||||  |||||  ::::
716 GlnSerAlaGlnIleSerLysAlaLeuValAlaGlyValaAspPheGlnAlaMetTrp 735
QY 2743 TATCCTCAGGAGACACAGATAGAGTTCTCTGAATCGGAGAACATTTAGTACGTCAT 2802
Db      |||  |||  |||  |||  |||  ::::  ::::  |||||
736 TyrThrAspGluAspHisGlyLeuAlaSerSerThrAlaHisGlnHisIleTyrSerHis 755
QY 2803 CTTTTCAGTCTTCAAGAA 2823
Db      ::::  |||||  |||||  ::::
756 MetSerHisPheLeuGlnGln 762

RESULT 9
A30107
dipeptidyl aminopeptidase B (EC 3.4.14.-) - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein YHR028c
C;Species: Saccharomyces cerevisiae
C;Date: 07-Jun-1990 #sequence_revision 30-May-1997 #text_change 09-Jul-2004
C;Accession: S46780; A30107
R;Du, Z.
submitted to the EMBL Data Library, June 1994
A;Description: The sequence of S. cerevisiae cosmid 8082.
A;Reference number: S46773
A;Accession: S46780
A;Molecule type: DNA
A;Residues: 1-818 <DUZ>
A;Cross-references: UNIPROT:PI18962; UNIPARC:UPI0000031A5F; EMBL:U10399; NID:g500689; PID
R;Roberts, C.J.; Pohlig, G.; Rothman, J.H.; Stevens, T.H.
J. Cell Biol. 108, 1363-1373, 1989
A;Title: Structure, biosynthesis, and localization of dipeptidyl aminopeptidase B, an in
A;Reference number: A30107; MUID:89174971; PMID:2647766
A;Accession: A30107
A;Molecule type: DNA
A;Residues: 1-82, 'H', 84-124, 'N', 126-181, 'LRRLET', 199-199, 'N', 201-365, 'DFKRGKERK', 376-57
A;Cross-references: UNIPARC:UPI0000172A3C; EMBL:X15484
A;Note: the authors translated the codon ACC for residue 572 as Asn
C;Genetics:
A;Gene: SGD:DAP2; SPE13; MIPS:YHR028c
A;Cross-references: SGD:S0001070, MIPS:YHR028c
A;Map position: 8R
C;Superfamily: dipeptidyl-peptidase IV
C;Keywords: dipeptidylpeptide hydrolase; glycoprotein; transmembrane protein; yeast vacu
F;30-45/Domain: transmembrane #status predicted <TM>
F;63,79,110,139,392,421/Binding site: carbohydrate (Asn) (covalent) #status predicted

Alignment Scores:
Pred. No.: 5,08e-27 Length: 818
Score: 484.50 Matches: 183
Percent Similarity: 39.4% Conservative: 95
Best Local Similarity: 26.0% Mismatches: 250
Query Match: 8.7% Indels: 177
DB: 1 Gaps: 35

US-10-825-632-2 (1-3120) x A30107 (1-818)
QY 847 ATTGCTTTTATACATAGCAACATATTTGGATATTAACATCGTAACACGAGAGAAAGG 906
Db      |||||  :::::  |||||  |||||  |||||  ::::  ::::  ::::  ::::  ::::  ::::
201 IleAlaTyrValGlnAspAsnAsnIleTyrIleTyrSerAlaIleSerLysLysThrIle 220
QY 907 AGACTCACTTATCTGCACAAATGAG-----CTAGCCAACATGGAGAAGATGCC 954
Db      |||  |||  |||  |||  |||  ::::  ::::  ::::  ::::  ::::  |||||
221 ArgAla-----ValThrAsnAspGlySerSerPheLeuPheAsnGlyLysProAsp--- 237
QY 955 AGATCAGCTGGAGTCGCTACCTTTGTTCTTCCAAAGAA-----TTTGATAGATATTTCTGGC 1011
Db      |||  |||  |||  |||  |||  ::::  ::::  |||||  |||||  ::::
238 -----TrpValTyrGluGluValPheValPheAspAspLysAla 250
QY 1012 TATTGGTGGTGTCAAAAGCTGAACAACTCCAGTGGTGTAATAATCTTAGAATCTTA 1071
Db      |||||  |||||  |||||  ::::
251 AlaTrpTrpSerProThrGlyAsp-----TyrLeuAla 261
QY 1072 TATGAAGAAATGATGAATCTCAGGTG---GAATATTATTCATGTTACATCCCTATGTTG 1128
Db      ::::  ::::  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
```


Tue Apr 18 08:18:36 2006

755 GluProArgAspIleThrGluValThrIlysPheIleGlnAenSerGlnHisIle 774
2398 GACTTAGATCGTGTGGGCATCCACGGCTGGTCTTATGGAGGATACCTCTCCCTGATGCCA 2457
775 AspLuserIysIleAlaIleTrpGlyTrpSertyrGlyGlyPheThrSerLeuIysThr 794
2458 TTA---ATGCAGAGGTTCAGATCTTCAGGTTGCTATTGCTGGGGCCCCAGTCACTCTG 2514
795 ValGluLeuAspAncGlyAspThrPheIysTyrAlaMetAlaValAlaIlePheValThrAen 814
2515 TGGATCTTCTATCATGATCAGAGTACACGGAAGTGTTATATGGCTCACCTGCACCAGAATAAA 2574
815 TrpThrLeuTyrAspSerValTyrThrGluArgTyrMetAenGlnProSerGluAenHis 834
2575 CAGGGCTATTACTTAGATCTTGGGCATGTCAGACGAAAAGTTCCCTCTGAACCAAAT 2634
835 GluGlyTyrPheGluValSerThrIleGlnAenPheIysSerPheGluSer---LeuLys 853
2635 CGTTTACTGCTCTTACATGGTTTTCTTGGAGAGAATGCCATTTTGGACATACAGTATA 2694
854 ArgLeuPheIleValHisGlyThrPheAspAspAenValHisIleGlnAenThrPheArg 873
2695 TTACTGAGTTTTTGTAGTGGGCTGGA---AAGCCATATGATTACAGATCTATCTCTCAG 2751
874 LeuValAspGlnLeuAenLeuLeuGlyLeuThrAsnTyrAspMetHisIlePheProAsp 893
2752 GAGAGACACACATTAAGAGTTCCTGAATCGGAGAACATATTGAATGTCATCTTTTGCAC 2811
894 SerAspHisSerIleArgTyrHisAenAlaGlnArgIleValPheGlnLysLeuTyrTyr 913
2812 TACTCTCAAGAAACCCTTGGATCAGCTATTGCTGCTCTAAAGTGATA 2859
914 TrpLeuArgAspAlaPheAlaGluArgPheAspAenThrGluValLeu 929

RESULT 11
T25174
hypothetical protein T23F1.7b - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T25174
R:Wilkinson, J.
submitted to the EMBL Data Library, October 1996
A:Reference number: Z19990
A:Status: preliminary; translated from GB/EMBL/DDBB
A:Molecule type: DNA
A:Residues: 1-799 <WIL>
A:Cross-references: UNIPROT:O18119; UNIPARC:UPI0000061190; EMBL:Z81129; PIDN:CAB03412.1.7.1
A:Experimental source: clone T23F1
C:Genetics:
A:Gene: CESP:T23F1.7b
A:Map position: 5
A:Incons: 13/3; 52/1; 111/3; 197/3; 366/3; 524/1; 557/1; 648/2; 703/1; 743/1
C:Superfamily: dipeptidyl-peptidase IV

Alignment Scores:

Pred. No.:	Length:
Score:	799
Percent Similarity:	Matches: 195
Best Local Similarity:	Conservative: 129
Query Match:	Mismatches: 298
DB:	Indels: 227
	Gaps: 36

US-10-825-632-2 (1-3120) x T25174 (1-799)

Qy 445 CCTCATTCAGACAGAACTATTACTCTGGCATGCT-----GGTGAGAAC 489
Dbb ||| :||| | ||| :||| :|||
Dbb 116 ProSerAlaAspArgLysTyrPheAlaMetMetAspHisAlaProAenProGlyMetAen 135
Qy 490 AGAGAAAAATACACTGTTTATTCTTGAATATCCC AAAATCATCATAGACGACGCTCTTA 549
Dbb ||| :||| | ||| :||| :|||
Dbb 136 ProGlnAenGluThrPheHisLeuIysIleValAenAenAenGluArgLeu----- 152

```
Db 391 AspThrLeuPhePheLeuLeuPro-----HisAspLysArg 402
QY 1609 ATTACATCTATTATTAAGGAAGCAAAATATAAAGCATCCAGTGGTGGCTGCTGCTCA 1668
Db 403 AspAsnAlaPheGlnValAlaSerLeuArgLeuSerHisGlyGlnLeuArgThrPro 422
QY 1669 AGTGATTTCAAGTCTCTATCAAGAGGAGATAGCAATTAACAGTGGTGAATGGGAAGTT 1728
Db 423 Lys-----PheLeuAsnLeuGlyGluArgAspVal 432
QY 1729 CTTGGCGCGCATGATCTAATATCAAGTTGATGATCGAGAGGCTGGTATATTGTGA 1788
Db 433 ThrSerIleAsnGlyIleAsnLysGluThr-----ArgThrIlePhePheHis 448
QY 1789 GGCACCAAGAGCTCCCTTTAGAGCATCACCTGATCGTAGTACGTAAATCTCTGGA 1848
Db 449 AlaAlaAlaProLysProSerHisArgSerLeuPheSerThrSer----- 463
QY 1849 GAGGTGACAGGCTGACGTGACGGTGTACTCAATCTTGTGCTGATCAGT----- 1899
Db 464 -----LeuAlaAspGluSerArgAsnSerAlaTyrCysIleSerCysSerIle 479
QY 1900 CAGCACTGTGACTTCTTTATAGTAAGTATAGTAAACCAAGAGATCCACACACTGTGTCTC 1959
Db 480 LysAsnCysThrTrpAlaGlnAlaGlnMetAspAspGlnMetLysThrAlaIleValSer 499
QY 1960 CTTTCAAGCTCAAGTCTCTGAAGAT-----GACCCAACCT 1995
Db 500 CysLysGlyProAlaAlaProHisThrAlaIleValAsnLeuThrArgMetAspSerAsp 519
QY 1996 TGCAAAACAAAG-----GAATTTTGGGCCCACTTTTGGATTCA 2034
Db 520 LysLysThrGluHisAlaAsnLeuLeuTyrAspLysThrTyrGlnAsnArgValGluGlu 539
QY 2035 GCAGTCTCTTCTGACTACTACTCTCCAGAAATTTCTTTTGAAGTACTACTGGA 2094
Db 540 AlaGly-----LeuProValIleIleLysGluThrIle-----LysIleSerAspAsp 555
QY 2095 TTTACATTTGTGGATGCTCTACAAGCTCATCATCTACAGCTCGGAAGAATAAT--- 2151
Db 556 PheAspAlaLeuIleLysLeuSerIleProLysAspIleTyrAsnArgAspLysHisGln 575
QY 2152 -----CCTACTGTCTCTTATATATGTTGGTCTCTCAGTGCAGTGGTGAATATCGG 2205
Db 576 AlaIleProLeuIleValHisValTyrGlyProAsnAspGln-----AsnThr 592
QY 2206 TTTAAGGAGTCAAGTATTTCCGCTTGAATACCTAGCTCTCTAGGTTATGGTTGTA 2265
Db 593 LysGluAlaThrGlnIleGlyIleGluValValAlaSerAlaSerGlnAlaAlaIle 612
QY 2266 GTG-----ATAGACACAGGGATCTGTGACGAGGCTTAATTTGAAGCGCTTAAA 2322
Db 613 LeuArgIleAspGlyArgGlySerGlyGlyArgGlyTrpLysTyrArgSerAlaIleTyr 632
QY 2323 TATAAATGGGTCAATAGAATTCAGCATCAGTGGAGCATCCCAATATCTAGTTCT 2382
Db 633 GlyGlnLeuGlyThrValGluValGluAspGlnIleLysAlaIleLysValValLeuArg 652
QY 2383 CGATAT---GATTTCAATTGACTTAGATCGTGGGCATCCACGGCTGGTCTTATGGAGA 2439
Db 653 LeuTyrArgHisLeuLeuAspAlaArgValAlaValPheGlyTrpSerTyrGlyGly 672
QY 2440 TACCTCTCCGTGATGGCATTAATGCAGAGTGCAGAT---ATCTTCAGGGTGGTATTGCT 2496
Db 673 PheMetThrLeuSerMetValAsnGluAlaProGluGlnPhePheLysCysAlaValSer 692
QY 2497 GGGGCCCCGAGTCACCTCTGTGATCTTCTATGATACAGATACACGGAACGTTATATGGGT 2556
Db 693 ValAlaProValThrAsnPheAlaTyrTyrAspAlaThrTyrThrGluArgTyrMetGly 712
QY 2557 CACCCCTGACCAAGTGAACGGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAG 2616
```

```
Db 713 -----AspAlaProLeuGluSerTyr-----SerAspValThrLysLysLeuAspAsn 728
QY 2617 TTCCCTCTGACCAAAATCGTTTACTGTCTTTACATGTTTCTCTGGATGAGAAATGTCAT 2676
Db 729 PheLysSer-----ThrArgLeuLeuMetHisGlyLeuLeuAspAsnValHis 746
QY 2677 TTTGCCACATACCATATATCTAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2736
Db 747 PheGlnAsnSerAlaIleLeuIleAspGluLeuGlnAsnArgGlyValAspPheAspLeu 766
QY 2737 CAGATCTATCTCCTCAGGAGACACACAGCATTAAGAGTTCCTGAATCGGAGAGATATGAA 2796
Db 767 MetValTyrProAsnGlnAlaHisSerLeuSerSerArgThrSer-----HisVal 784
QY 2797 CTGCATCTTTTTCACCTACCTTCAAGAA 2823
Db 785 GlyLysMetThrHisPheLeuArgGln 793
RESULT 12
T25173
hypothetical protein T23F1.7a - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T25173
R:Wilkinson, J.
submitted to the EMBL Data Library, October 1996
A:Reference number: Z19990
A:Accession: T25173
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-779 <WIL>
A:Cross-references: UNIPROT:O18119; UNIPARC:UPI000002A220; EMBL:Z81129; PIDN:CAB03411.1
A:Experimental source: clone T23F1
C:Genetics:
A:Gene: CESP:T23F1.7a
A:Map position: 5
A:Introns: 13/3; 52/1; 111/3; 151/2; 177/3; 346/3; 504/1; 537/1; 628/2; 683/1; 723/1
C:Superfamily: dipeptidyl-peptidase IV
Alignment Scores:
Pred. No.: 3,39e-24 Length: 779
Score: 445.50 Matches: 193
Percent Similarity: 38.0% Conservative: 126
Best Local Similarity: 23.0% Mismatches: 292
Query Match: 8.0% Indels: 229
DB: 2 Gaps: 38
US-10-825-632-2 (1-3120) x T25173 (1-779)
QY 445 CCTCATTCAGACAGATCTATTACCTTGCCATGCTCT-----GGTCAGAAC 489
Db 116 ProSerAlaAspArgLysTyrPheAlaMetMetAspHisAlaProAsnProGlyMetAsn 135
QY 490 AGAGAAAATACACTGTTTATTCTGAAATTCCTCAAACTATCAATAGACAGCAGTCTTA 549
Db 136 ProGlnAsnGluThrPheHisLeuLysIleValAsnAsnAsnGluArg----- 151
QY 550 ATGCTCTCTTGGAGGCTCTTTTGGATCTTTTTCAGGCAACACTGCGATGGAATGTAT 609
Db 152 -----IleThrTyrAspIleGlyLeu--- 158
QY 610 TCTCGAAGAAGAAGACTATTAAAGAGAAAGAAACGCAATTGGAACAGTCGGAATGCTTCT 669
Db 159 ---ArgLysGluGluSerVal-----IleGlnAla 167
QY 670 TAGCATTTATCAACAAGAAAGTGGAAACATTTCTGTGTTTCAAGCCGGTAGTGGAAATTTAC 729
Db 168 PheLysTrpAsnGlyLysPheAsnAspPheValPheValGluSerAsnLysIleTyrTyr 187
QY 730 GTAAAAGATGGAGGGCCACAGAGATTTCAGCACACACCTTTAAGGCCCAATCTAGTGAA 789
Db 187 ----- 187
```


dipeptidyl aminopeptidase like protein - human

C:Species: Homo sapiens (man)

C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004

C:Accession: I54331

R:Yokotani, N.; Doi, K.; Wentholt, R.J.; Wada, K.

Hum. Mol. Genet. 2, 1037-1039, 1993

A:Title: Non-conservation of a catalytic residue in a dipeptidyl aminopeptidase IV-related

A:Reference number: I54331; MUID:93372805; PMID:8103397

A:Accession: I54331

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-865 <RES>

A:Cross-references: UNIPROT:P42658; UNIPARC:UPI0000047378; GB:M96859; NID:g306705; PIDN:

C:Superfamily: dipeptidyl-peptidase IV

Alignment Scores:

Pred. No.:	4.45e-24	Length:	865
Score:	444.00	Matches:	222
Percent Similarity:	37.0%	Conservative:	145
Best Local Similarity:	22.4%	Mismatches:	385
Query Match:	8.0%	Indels:	239
DB:	2	Gaps:	40

US-10-825-632-2 (1-3120) x I54331 (1-865)

QY	82	CGGCGCGCGCGCGAGGAAGCCACTGC-----AACACAGGACCGGAGTGGAGGCGG	132
DB	19	ProAlaProPro-GluAlaSerHisLeuLeuGlyGlyGlyProGluGluAAspGlyG1	38
QY	133	CGACAGATGAAGCGCGCGCGCGTCCATAGCGCA-----CGTCGGGA	177
DB	38	yAlaGlyAlaLysProLeuGlyProArgAlaGlnAlaAlaProArgGluAAspGlyG1	58
QY	178	CGTCCGCGCGCGCGCGCGGGAAGGAATGCAACATGG-----	217
DB	58	yGlyGlyGlyAlaGlyArgProArgPheGlnTyrGlnGlyArgSerAspGlyAs	78
QY	218	-----CAGCAGCAATCGGAACAGACAGCTG	243
DB	78	pGluGluAspGluLeuValGlySerAsnProProGlnArgAsnTrpLys	94
QY	244	GGTGTGAGATATTGAACTGCGGACTGTGAGGAGAAATATGAATCAGAGTCCGCT	303
DB	95	-----GlyI1	96
QY	304	AAATTGGAGCCTTTTATGTTGAGCGGTATCTCTGAGTCAGCTTAAAGAGCTGCTG--	361
DB	96	eAlaIleAlaLeuValIleLeuValIleCysSerLeuIleValThrSerValIleLe	116
QY	362	-----CCGATACCAAAATATCATGCTCATGCTTAAGCTTAAGCCACCATGATTC	414
DB	116	uLeuThrProAlaGluAspAsn-----	123
QY	415	ATGTTGTGAAGAGGAATGATCCAGATGACCTCATTCAGACAGAACTTATACCTTGCC	474
DB	124	-----SerLeuSerGlnLysLysValThrValG1	134
QY	475	ATGTCGTGGTGAACACAGAG---AAATACACTGTTTATCTGAAATTCCTCAAAACTATC	531
DB	134	uAspLeuPheSerGluAspPheLysIleHis-----As	145
QY	532	AATAGACAGCAGCTTAAATGCTCTCTGGAAGCCTCTTTTGGATCTTTTTCAGGCAACA	591
DB	145	pProGlu-AlaLysTrpIleSerAspThrGluPheIleTyrArgGluGlnLysGlyThrV	165
QY	592	CTGACTATGGAATGATTTCTCGAAGAGAAAGAACTATTAAAG---GAAAGAAACGCA	648
DB	165	alArgLeuTrpAsnValGluThrAsnThrSerThrVal-LeuIleGluGlyLysIle	184
QY	649	GGAACAGTCCGAATTTGCTTCTTACGATTATACCAAGGAAGTGGAACTTTCTGTTT---	705
DB	185	GluSerLeuArgAlaIleArgTyrGluIleSerProAspArgGluTyrAlaLeuPheSer	204

QY	706	-----CAAGCGGTAGTGGAAATTTATCAGCTAAAGATGGAGG	744
DB	205	TyrAsnValGluProIleTyrGlnHisSerTyrThrGlyTyrValLeuSerLysIle	224
QY	745	CCACAGAGATTACGCAACACTTTAAGGCCCAATCTAGTGGAACTAGTTGTCCCAAC	804
DB	225	ProHisGlyAspProGlnSerLeuAAspProGluValSerAsnAlaLysLeuGlnTyr	244
QY	805	ATACGGATGGATCCAAAATTTATGCCCGCTGATCCAGACTGGATTGCTTTTATACATAGC	864
DB	245	AlaGlyTrpGlyProLys-----GlyGlnGlnLeuIlePheIleGlu	259
QY	865	AACGATATTTGGATATCTAACATCGTAACAGAGAAAGAGGAGACTCACTTATGTGCAC	924
DB	260	AsnAsnIleTyrTyrCysAlaHisValGlyLysGlnAlaIleArgVal-----	275
QY	925	AATGAGCTAGCCACATGGAAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTC	984
DB	276	-----ValSerThrGlyLysGluGlyValIleTyrAsnGlyLeuSerAspTrpLeuTyr	293
QY	985	CAAGAAGAA---TTTGATAGATATCTCTGCTATTGCTGTGTCCTCAAAAGCTGAAACAACT	1041
DB	294	GluGluGluLeuLeuLysThrHisIleAlaHisTTrpTrpSerProAspGly-----	310
QY	1042	CCAGTGGTGTAAATTTCTAGATTTCTATAGAATTTATAGAATAATGATCTGAGTGGAA	1101
DB	311	-----ThrArgLeuAlaTyrAlaAlaIleAsnAspSerArgValPro	324
QY	1102	ATTATTCATGTT-----ACATCCCTATGTTGGAAACAGAGGGCGAGATCATTC	1152
DB	325	IleMetGluLeuProThrTyrThrGlySerIleTyrProThr-----ValLysProTyr	342
QY	1153	CGTTATCTTAAACAGGTACAGCAATCTTAAGTCATCTTTTAAGATGTCAGAAATATG	1212
DB	343	HisTyrProLysAlaGlySerGluAsnProSerIleSerLeuHis-----Val	358
QY	1213	ATTGATCTGAGGAAGGATCATAGATGTCATAGATAGGAACATAATCAACCTTTTCAG	1272
DB	359	IleGlyLeuAsnGlyProThrHisAspLeu-----GluMetProProAspAsp	375
QY	1273	ATTCTATTGAAGAGTTGAATATATATGTCAGAGCTGGAGTCTCTGAGGAGAAATAT	1332
DB	376	ProArgMetArgGluTyr---TyrIleThrMetValLysTrpAlaThrSerThrLysVal	394
QY	1333	GCT-----TGTTCATCTCTAGATCGCTCCAGACTCGCTACAGATAGTGTGATC	1386
DB	395	AlaValThrTrp-----LeuAsnArgAlaGlnAsn-----Val	405
QY	1387	TCACCTGAATTTATTCACAGTAGAAGATGATTTATGGAAAGCAGACAGACTCATTTAG	1446
DB	406	Ser-----IleLeu	408
QY	1447	TCAGTGCCTGATCTGTGAGCCCACTAATTATC-----TATCAAGAAACACACATC	1500
DB	409	ThrLeuCysAspAlaThrThrGlyValCysThrLysHisGluAspGluSerGluAla	428
QY	1501	TGGATAATATCCATGACATCTTTTCATGTTTTCCTCCCAAGTACAGAGAGGAAAT---	1557
DB	429	TrpLeuHis-----ArgGlnAsnGluProValPhe	439
QY	1558	-----GAGTTTATTTTTCCTGCTGAATGCAAAACAGGTTTCCTG---CAT	1599
DB	440	SerLysAspGlyArgLysPhePheIleArgAlaIleProGlnGlyGlyArgGlyLys	459
QY	1600	TTATACAAAATTACATCTATTTTAAAGGAAACAAATAATAACCATCCAGTGGTGGCTG	1659
DB	460	PheTyrHisIleThr-----ValSerSerSerGlnProAsnSerSerAsnAspAsnIle	477
QY	1660	CCTGCTCCAAGTATTTCAGTGTCTTATCAAAAGAGGAGATAGCAATTCACAGTGGTGA	1719
DB	478	Gln-----SerIleThrSerGlyAsp	484
QY	1720	TGGAGAGTCTTGGCGGCGATGATCTAATATCAAGTTGATGAAGTCAGAGGCTGGTA	1779

		:::	:::	307	GlumMetThrProAspAspProArgMetArgGluTyr---TyrIleThrMetValLys	325
		:::	:::	1312	TGGACTCCTCGAGGGAAATATGCT-----TGTCTCATCTACTAGATCGCTCCACGACT	1365
		:::	:::	326	TpAlaIthrSerThrLysValAlaValAsnTrp-----LeuSerArgAlaGlnAsn	342
		:::	:::	1366	CGCTACAGATAGTGTGTGATCTCACCTACCTGAATATTATTCCTCCAGTAGAAGATGATGTTATG	1425
		:::	:::	343	-----ValSer-----	344
		:::	:::	1426	GAAGGCAGAGACTCATTTGAGTCACTGCTCGCTGATTCTGTGACGCCACTAAATTATC-----	1479
		:::	:::	345	-----IleLeuThrLeuCysAspAlaThrThrGlyValCysThrLysLys	359
		:::	:::	1480	TATGAAGAAACAAACAGACATCTCGATATAAATATCCATGACATCTTTTCATGTTTTCCTCCAA	1539
		:::	:::	360	HisGluAspGluSerGluAlaTrpLeu-----	368
		:::	:::	1540	AGTCACGAGAGGAAATTGAGTATTATTTTGGCTCTGAAATGCAAAACAGGTTTCGGTCAT	1599
		:::	:::	369	---HisArgGlnAsnGluGluProValPheSer-----LysAspGly---ArgLys	383
		:::	:::	1600	TTATACAAATTCATCTATTTTAAAG-----GAAAGCAAAATATAACGATCCAGTCGT	1653
		:::	:::	384	PhePhePheValArgAlaIleProGlnGlyGlyGlnGlyLysPheTyrHisIleThrVal	403
		:::	:::	1654	GGCTGCCTGCTCCAAGTGATTTCACGTGCTCTATCAAGAGAGAGATAGCAATACCAAGT	1713
		:::	:::	404	SerSerSerGlnProAsnSerSerAsnAspAsnIleGln-----SerIleThrSer	420
		:::	:::	1714	GGTGAATGGGAAGTTCTTGGCCGCATGGATCTAATATCCAAAGTTGATGAAGTACAGAAGG	1773
		:::	:::	421	GlyAspTrpAspValThr-----LysIleLeuSerTyrAspGluLysArgSer	436
		:::	:::	1774	CTGGTATATTTGAAGGCACCAAGACTCCCTTTAGACATCACTCTAGCTAGTCAGT	1833
		:::	:::	437	GlnIleTyrPheLeuSerThrGluAspLeuProArgArgGlnLeuTyrSerAlaSer	456
		:::	:::	1834	TACGTAAATCTCGAGAGGTGCACAGG-----CTGACTGACCGGTGGC	1875
		:::	:::	457	ThrVal-----GlySerPheAsnArgGlnCysLeuSerCysAspLeuValAspAsnCys	474
		:::	:::	1876	TACTCACATTTCTGTGCATCAGTCAGCTAGTACTCTTTATAAGTAAGTATAGTAAC	1935
		:::	:::	475	ThrTyrPheSerAlaSerPheSerProGlyAlaAspPhePheLeuLeuLysCysGluGly	494
		:::	:::	1936	CAGAAGATCCACACTGTGTGTCTCTTTACAGCTATCAAGTCTCTGGAAGATGACCCCACT	1995
		:::	:::	495	ProGlyValPro---ThrValSerValHisAsnThrThrAspLysLysLysMetPheAsp	513
		:::	:::	1996	TGCAAAACAAAGGAA---TTTTGGCCACCATTTTCGATTTCAGCAGGTCCTCTTCCTGCAC	2052
		:::	:::	514	LeuGluThrAsnGluHisValGlnLysAlaIleSerAspArgGlnMetProLysValGlu	533
		:::	:::	2053	TATACTCTCCAGAAATTTCTCTTTTGAAAGTACTACTCGGATTTTACATCTGTATGGGATG	2112
		:::	:::	534	TyrArgLysIleGlu-----ThrAspAspTyrAsnLeuProIleGln	547
		:::	:::	2113	CTCTACAAAGCCTCATGATCTACAGCCTGGAAAGAAATATCTCTACTGTGCTGTTCATATAT	2172
		:::	:::	548	IleLeuLysProAlaThrPheThrAspThrAlaHisTyrProLeuLeuLeuValValAsp	567
		:::	:::	2173	GGTGTCTCAGTTCGAGTGGTGAATAATTCGGGTTTAAAGGAGTCAAGTATTTCCTCGCTTG	2232
		:::	:::	568	GlyThrProGlySerGlnSerValAlaGluLysPhe-----AlaValThrTrpGluThr	585
		:::	:::	2233	AATACCTAGCCTCTCAGGTATGTGGTTGTAGTGATAGACAAACAGGGGATCTCTGTCCAC	2292
		:::	:::	586	ValMetValSerSerHisGlyAlaValValValLysCysAspGlyArgGlySerGlyPhe	605
		:::	:::	2293	CGAGGGCTTAAATTTGAAGGCGCCTTTTAAATATATAAATGGGTCAAATGAAATTGACGAT	2352

Db	606	GlnGlyThrArgLeuLeuHisGluValArgArgLeuGlySerLeuGluGluValAsp	625
Qy	2353	CAGGTGGAAGCACTCCAATATCTTCGATATGATTTCATTGACATTAGATCGTGTG	2412
Db	626	GlnMetGluAlaValArgValMetLeu---LysGluProTyrIleAspLysThrArgVal	644
Qy	2413	GGCATCCAGCGGTGCTTCATGAGAGATACCTCTCCCTGATGGCATTAATCGAGAGTCA	2472
Db	645	AlaValPheGlyLeAspTyrGlyGlyTyrLeuSerThrTyrLeuLeuProAlaLysGly	664
Qy	2473	GAT-----ATCTTCAGGGTTGCTATTGCTGGGGCCCGCCAGTCACTCTGTGGATC	2520
Db	665	AspGlyGlnAlaProValPheSerCysGlySerAlaLeuSerProIleThrAspPheLys	684
Qy	2521	TTCTATGATACAGGATACACGGAAGCTTATATGGGTCCACCTCGACAGAATGAACAGGCG	2580
Db	685	LeuTyrAlaSerAlaPheSerGluArgTyrLeuGlyLeuHisGlyLeuAspAsnArgAla	704
Qy	2581	TATTACTTAGGATCTGTGGCCATCGACGAAGAAAGTCCCTCTGAACCAAAATCGTTTA	2640
Db	705	TyrGluMetAlaLysValAlaHisArgValSerAlaLeu-----GluGlyGlnGlnPhe	722
Qy	2641	CTGCTCTTACATGGTTCTTCGGATGAGAAATGTCATTTTGCATACACAGTATATTACTG	2700
Db	723	LeuValIleHisAlaThrAlaAspGluLysIleHisPheGlnHisThrAlaGluLeuIle	742
Qy	2701	AGTTTTTATGTCAGGCGTGGAAAGCCATATGATTACAGATCTATCTCTCAGAGAGACAC	2760
Db	743	ThrGlnLeuIleLysGlyLeAlaAsnTyrSerLeuGlnIleTyrProAspGluSerHis	762
Qy	2761	AGCATAGAAGTTCCTGAATCGGGGAACATATTATGAATCGATCTTTTTCGACTACCTCAA	2820
Db	763	TyrPheSerSerAlaAlaLeuGlnGlnHisLeuHisArgSerIleLeuGlyPhePheVal	782
Qy	2821	GAA 2823	
Db	783	Glu 783	
RESULT 15			
I68600			
dipeptidyl aminopeptidase like protein - human			
C/Species: Homo sapiens (man)			
C/Date: 06-Sep-1996 #sequence_revision 06-Sep-1996 #text_change 21-Jul-2000			
C/Accession: I68600			
R/Yokotani, N.; Doi, K.; Wenthhold, R.J.; Wada, K.			
Hum. Mol. Genet. 2, 1037-1039, 1993			
A/Title: Non-conservation of a catalytic residue in a dipeptidyl aminopeptidase			
A/Reference number: I54331; MUD:93372805; PMID:8103397			
A/Accession: I68600			
A/Status: preliminary;			
A/Molecule type: mRNA			
A/Residues: 1-803 <RES>			
A/Cross-references: UNIPARC:UPI000016A7F9; GB:M96860; NID:G306707; PIDN:AAA35756			
C/Superfamily: dipeptidyl-peptidase IV			
Alignment Scores:			
Pred. No.:	2,99e-23	Length:	803
Score:	432.50	Matches:	185
Percent Similarity:	39.4%	Conservative:	127
Best Local Similarity:	23.4%	Mismatches:	326
Query Match:	7.8%	Indels:	154
DB:	2	Gaps:	32
US-10-825-632-2 (1-3120) x I68600 (1-803)			
Qy	585	GGCACTGCGACTGATGGAATGATTCCTCGAGAAGAAGAACTATTAAAGAA	641
Db	101	GlyThrValArgLeuTrpAsnValGluThrAsnThrSerThrVal-LeuIleGluGlyLy	120
Qy	642	ACGCATTGGAAACAGTCGGAATTGCTTCTACGATTATCACCAAGGAAGTGAACATTCT	701
Db	120	sysIleGluSerLeuArgAlaIleArgIleGluIleSerProAspArgGluTyrAlaLe	140

Db 752 rSerLeuGlnIleTyrProAspGluSerHisTyrPheThrSerSerSerLeulysGlnHi 772
QY 2790 TTATGAACTGCATCTTTTGCACTACCTTCAAGAA 2823
Db 772 sLeuTyrArgSerIleIleAsnPhePheValGlu 783

Search completed: April 14, 2006, 11:34:39
Job time : 195.5 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: April 14, 2006, 10:51:36 ; Search time 137.6 Seconds
(without alignments)
3199.489 Million cell updates/sec

Title: US-10-825-632-2
Perfect score: 5552
Sequence: 1 aagtgtctaaagctccgagg.....agaattactataaaaaaaa 3120

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 4332886

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xl-p
-O=/abs/ABSSWEB.spool/US10825632/runat_14042006_092349_3432/app_query.fasta_1
-DB=UniProt -QMT=fastan -SUFFIX=n2p.rup -MINMATCH=0.1 -LOOPECL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss02p
-USER=US10825632 @CGN 1 1 466 @runat_14042006_092349_3432 -NCP=6 -ICPU=3
-NO MWAP -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-XGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : UniProt_05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4726	85.1	898	1 DPP8 HUMAN	Q6vix1 homo sapien
2	4559.5	82.1	892	1 DPP8 MOUSE	Q80ya7 mus musculus
3	2870	51.7	863	1 DPP9 HUMAN	Q86ti2 homo sapien
4	2833	51.0	862	1 DPP9 MOUSE	Q8bvg4 mus musculus
5	2829.5	51.0	923	2 Q4SM6 TETNG	Q4sm6 tetraodon n
6	2707.5	48.8	847	2 Q6GR22 XENLA	Q6gr22 xenopus lae
7	1643.5	29.6	508	2 Q75273 HUMAN	Q75273 homo sapien
8	1641	29.6	886	2 Q7QBK1 ANOGA	Q7qbki anopheles g
9	1624.5	29.3	1053	2 Q9VC20 DROME	Q9vc20 drosophila
10	1607.5	29.0	1113	2 Q9VC19 DROME	Q9vc19 drosophila
11	1567	28.2	740	2 Q5TK8 ANOGA	Q5tk8 anopheles g
12	1321	23.8	621	2 Q7PTT8 ANOGA	Q7ptt8 anopheles g
13	1142.5	20.6	557	2 Q5TXJ2 ANOGA	Q5txj2 anopheles g
14	1124.5	20.3	803	2 Q54U01 DICDI	Q54u01 dictyosteli
15	1049	18.9	432	2 Q75868 HUMAN	Q75868 homo sapien
16	924.5	16.7	927	2 Q965K3_CAEL	Q965k3 caenorhabdi

17	922.5	16.6	931	2	O44987_CAEL	O44987 caenorhabdi
18	867.5	15.6	895	2	O61CU7_SHEBR	O61cu7 caenorhabdi
19	846	15.2	763	2	Q8EAB7_SHEON	Q8eab7 shewanella
20	842	15.2	738	2	O9A6B0_CAUCR	O9a6b0 caulobacter
21	837.5	15.1	746	2	O9FNF6_ARATH	O9fnf6 arabidopsis
22	792.5	14.3	745	2	O6F3I7_PSED	O6f3i7 pseudomonas
23	786	14.2	596	2	O6K8B0_ORISA	O6k8b0 oryza sativ
24	779	14.0	743	2	O5QX36_IDILO	O5qx36 idiomarina
25	759	13.7	745	2	O5H5W8_XANOR	O5h5w8 xanthomonas
26	754	13.6	741	2	P95782_XANWA	P95782 xanthomonas
27	751.5	13.5	751	2	O4UPD3_XANCP	O4upd3 xanthomonas
28	751.5	13.5	751	2	O8P3V8_XANCP	O8p3v8 xanthomonas
29	746	13.4	766	2	O4TNE1_SPHEN	O4tnel erythrobact
30	744.5	13.4	757	2	O8PFD7_XANAC	O8pfd7 xanthomonas
31	727	13.1	552	2	O8GUJ7_ARATH	O8guj7 arabidopsis
32	724.5	13.0	720	2	O5NNM8_ZYMMO	O5nnm8 zymomonas m
33	722.5	13.0	749	2	O7NEK8_GLOVI	O7nek8 gloebacter
34	672	12.1	736	2	O8A028_BACTN	O8a028 bacteroides
35	670	12.1	739	2	O5FQI6_GLUOX	O5fqy6 gluconobact
36	655.5	11.8	850	2	O6H9E3_3TRYP	O6h9e3 trypanosoma
37	651	11.7	736	2	O5LGU5_BACFN	O5lgu5 bacteroides
38	648	11.7	736	2	O64XP9_BACFR	O64xp9 bacteroides
39	643	11.6	852	2	O4QIH9_LEIMA	O4qih9 leishmania
40	620	11.2	711	2	O47900_FIAME	O47900 flavobacter
41	606	10.9	809	2	O4NVV1_9DELT	O4nvv1 anaeromyxob
42	595.5	10.7	778	2	O51ZF7_MAGGR	O51zf7 magnaporthe
43	587	10.6	769	2	O5QVY7_IDILO	O5qvyy idiomarina
44	582.5	10.5	771	2	O42812_ASPOR	O42812 aspergillus
45	578.5	10.4	731	2	O6L872_PREIN	O6l872 prevotella

ALIGNMENTS

RESULT 1
DPP8_HUMAN . STANDARD; PRT; 898 AA.
ID DPP8_HUMAN . STANDARD; Q7Z4D3; Q7Z4E1; Q8IMG7; Q8NEM5; Q96JX1; Q9HBM2;
AC Q6VIX1; Q7Z4C9; Q7Z4D3; Q7Z4E1; Q8IMG7; Q8NEM5; Q9HBM2;
AC Q9HBM3; Q9HBM4; Q9HBM5; Q9NKF4;
DT 13-SEP-2005 (Rel. 48, Created)
DT 13-SEP-2005 (Rel. 48, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Dipeptidyl peptidase 8 (EC 3.4.14.5) (Dipeptidyl peptidase VIII) (DPP8)
DE (Prolyl) dipeptidase DPP8 (Dipeptidyl peptidase IV-related protein 1)
DE (DPP-1)
GN Name=DPP8; Synonyms=DPP1; ORFNames=MSTP097, MSTP135, MSTP141;
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 3), NUCLEOTIDE SEQUENCE [MRNA] OF 334-898 (ISOFORM 4), NUCLEOTIDE SEQUENCE [MRNA] OF 540-898 (ISOFORM 5), NUCLEOTIDE SEQUENCE [MRNA] OF 260-792 (ISOFORM 6), FUNCTION, CATALYTIC ACTIVITY, ENZYME REGULATION, TISSUE SPECIFICITY, INDUCTION, AND SUBCELLULAR LOCATION.
RC TISSUE=Placenta;
RX MEDLINE=20467194; PubMed=11012666;
RA Abbott C.A., Yu D.M.T., Woollatt E., Sutherland G.R., McCaughan G.W., Gorrell M.D.;
RA "Cloning, expression and chromosomal localization of a novel human dipeptidyl peptidase (DPP) IV homolog, DPP8.";
PL Eur. J. Biochem. 267:6140-6150(2000).
RN [2]
RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 3), CATALYTIC ACTIVITY, ENZYME REGULATION, BIOPHYSICOCHEMICAL PROPERTIES, SUBCELLULAR LOCATION, AND TISSUE=Testis;
RX PubMed=12662155; DOI=10.1042/BJ20021914;
RA Qi S.Y., Riviere P.J., Trojanar J., Junien J.-L., Akinsanya K.O.;
RA "Cloning and characterization of dipeptidyl peptidase 10, a new member of an emerging subgroup of serine proteases.";
RT Biochem. J. 373:179-189(2003).

[3] NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1).
 RP TISSUE=Testis;
 RA Sha J.H., Zhou Z.M., Li J.M.;
 RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.

[4] NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 4), AND NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 211-898 (ISOFORM 2).
 RP TISSUE=Hepatoma, and Placenta;
 RX PubMed=14702039; DOI=10.1038/ng1285;
 RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R., Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H., Sekine M., Ohyashi M., Nishi T., Shibahara T., Tanaka T., Ishii S., Yamamoto J.-I., Saito K., Kawai T., Iwayanagi T., Wagatsuma M., Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa K., Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K., Tanai K., Kimata M., Watanabe M., Hirakawa S., Chiba Y., Ishida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T., Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O., Nomura Y., Togai S., Komai F., Hara R., Takeuchi K., Arita M., Imose N., Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S., Yoshioka K., Matsunawa H., Ichihara T., Shichida N., Sano S., Morya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O., Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B., Yamazaki M., Koniya M., Tashiro H., Tanigami A., Fujiwara T., Fujitani Y., Yamada K., Fujii Y., Ozaki K., Hirao M., Omori Y., Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S., Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T., Takamura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M., Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T., Nakagawa K., Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Yamashita R., Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R., Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
 RT "Complete sequencing and characterization of 21,243 full-length human cDNAs.";
 RL Nat. Genet. 36:40-45(2004).

[5] NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORMS 1 AND 2).
 RP TISSUE=Testis;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Uginis T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Lochell N.A., Peters G.J., Abramson R.D., Mullaly S.J., Richardson J., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Rodriguez A.C., Kravitz M.I., Skalska U., Smailus D.E., Butlerfield Y.S.N., Kravitz M.I., Skalska U., Smailus D.E., Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16999-16903(2002).

[6] NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 561-898.
 RP TISSUE=Aorta;
 RA Zhao B., Xu H.S., Tong Y.K., Sheng H., Qin B.M., Liu Y.Q., Liu B., Wang X., Zhang Q., Song L., Gao Y., Zhang C.L., Ye J., Ji X.J., Liu B.H., Lu H., Chen J.Z., Cai M.Q., Zheng W.Y., Teng C.Y., Liu Q., Yu L.T., Lin J., Gong Q., Zhang A.M., Gao R.L., Hui R.T.;
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.

[7] MUTAGENESIS OF GLU-275; SER-755; ASP-833 AND HIS-865, CATALYTIC ACTIVITY, BIOPHYSICOCHEMICAL PROPERTIES, AND SUBCELLULAR LOCATION.
 RP PubMed=12534281; DOI=10.1021/bi026946s;
 RX Ajami K., Abbott C.A., Obradovic M., Gysbers V., Kaehne T., McCaughan G.W., Gorrell M.D.;
 RA "Structural requirements for catalysis, expression, and dimerization in the CD26/bpIV gene family.";
 RL Biochemistry 42:694-701(2003).

[8] BIOPHYSICOCHEMICAL PROPERTIES.
 RP PubMed=15039077; DOI=10.1016/j.bmcl.2004.11.023;
 RX Chen Y.-S., Chien C.-H., Goparaju C.M., Hsu J.T.-A., Liang P.-H., Chen X.;
 RA "Purification and characterization of human prolyl dipeptidase DPP8 in Sf9 insect cells.";
 RL Protein Expr. Purif. 35:142-146(2004).

[9] ENZYME REGULATION.
 RP PubMed=15664838; DOI=10.1016/j.bmcl.2004.11.023;
 RX Jiang W.-T., Chen Y.-S., Hsu T., Wu S.-H., Chien C.-H., Chang C.-N., Chang S.-P., Lee S.-J., Chen X.;
 RA "Novel isoindoline compounds for potent and selective inhibition of prolyl dipeptidase DPP8.";
 RL Bioorg. Med. Chem. Lett. 15:687-691(2005).

CC -1- FUNCTION: Dipeptidyl peptidase that cleaves off N-terminal dipeptides from proteins having a Pro or Ala residue at position 2. May play a role in T-cell activation and immune function. Yaa-1-2. CATALYTIC ACTIVITY: Release of an N-terminal dipeptide, Xaa-Yaa-|-Zaa, from a polypeptide, preferentially when Yaa is Pro, provided Zaa is neither Pro nor hydroxyproline.
 CC -1- ENZYME REGULATION: Inhibited by zinc. Inhibited by the serine protease inhibitor 4-(2-aminoethyl)benzenesulphonyl fluoride (AEPF), and by di-isopropylfluorophosphate. Specifically inhibited by isoindoline derivatives.
 CC -1- BIOPHYSICOCHEMICAL PROPERTIES:
 CC Kinetic parameters:
 CC KM=208 uM for Ala-Pro-AMC;
 CC KM=130 uM for Ala-Pro-AFC;
 CC KM=120 uM for H-Ala-Pro-pNa;
 CC KM=1420 uM for H-Ala-Ala-pNa;
 CC KM=310 uM for H-Arg-Pro-pNa;
 CC KM=2050 uM for H-Asp-Pro-pNa;
 CC KM=480 uM for H-Gly-Pro-pNa;
 CC pH dependence:
 CC Optimum pH is 7.4-8.5. Little activity below pH 6.5;
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=6;
 CC Name=1;
 CC IsoId=Q6V1X1-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=Q6V1X1-2; Sequence=VSP_013864;
 CC Name=3;
 CC IsoId=Q6V1X1-3; Sequence=VSP_013860;
 CC Name=4;
 CC IsoId=Q6V1X1-4; Sequence=VSP_013860, VSP_013862;
 CC Name=5;
 CC IsoId=Q6V1X1-5; Sequence=VSP_013863;
 CC Name=6;
 CC IsoId=Q6V1X1-6; Sequence=VSP_013861;
 CC -1- TISSUE SPECIFICITY: Ubiquitously expressed, with highest levels in testis, placenta, prostate, muscle and brain.
 CC -1- INDUCTION: In activated T-cells.
 CC -1- SIMILARITY: Belongs to the peptidase S9B family. DPPIV subfamily.
 CC -1- CAUTION: Ref.4 (BAA91059) sequence differs from that shown due to a frameshift in position 486.
 CC -1- CAUTION: Ref.5 (AAQ13650 and AAQ13623) sequences differ from that shown due to several frameshifts.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its

use as long as its content is in no way modified and this statement is not removed.

CC
CC
CC
DR EMBL; AF221634; AAG29766.1; -; mRNA.
DR EMBL; AF221635; AAG29767.1; -; mRNA.
DR EMBL; AF221636; AAG29768.1; -; mRNA.
DR EMBL; AF221637; AAG29769.1; -; mRNA.
DR EMBL; AF172659; AA017261.1; -; mRNA.
DR EMBL; AY354202; AAQ63887.1; -; mRNA.
DR EMBL; AK000290; BAA91059.1; ALT_FRAME; mRNA.
DR EMBL; AK027826; BAB55395.1; ALT_INIT; mRNA.
DR EMBL; BC030688; AAH30688.3; -; mRNA.
DR EMBL; BC040203; AAH40203.1; ALT_INIT; mRNA.
DR EMBL; AF176779; AAQ13657.1; ALT_INIT; mRNA.
DR EMBL; AF175225; AAQ13650.1; ALT_FRAME; mRNA.
DR EMBL; AF173382; AAQ13623.1; ALT_FRAME; mRNA.

Alignment Scores:
Pred. No.: 0 Length: 898
Score: 4726.00 Matches: 886
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 85.1% Indels: 0
DB: 1 Gaps: 0

US-10-825-632-2 (1-3120) x DPP8_HUMAN (1-898)

QY 202 GGAAATGCAATGGCAGCAGCAATGGAACAGACAGCTGGTGTGAGATTTTGA 261
DB 13 GlyLysCysAsnMetAlaAlaMetGluThrGluGlnLeuGlyValGluLeuPheGlu 32
QY 262 ACTGGCGACTGTGAGGAGAATTTGAATCACAGATCGGCTAAATGGAGCCTTTTAT 321
DB 33 ThrAlaAspCysGluGlnAsnIleGluSerGlnAspArgProLysLeuGluProPheTyr 52
QY 322 GTTGAGCGGTATTCCTGGAGTCAGCTTAAAGAGTCGTCCGATACCCAGAAATATCAT 381
DB 53 ValGluArgTyrSerTrpSerGlnLeuLysLeuLeuAlaAspThrArgLysTyrHis 72
QY 382 GGCTACATGATCGCTAAGGCCACCATGATTTTCATGTTGTGAGAGGAATGATCAGAT 441
DB 73 GlyTyrMetMetAlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAsp 92
QY 442 GGACCTCATTCAGACAGAAATCTATTACCTTCCGATGTCTGGTGAACAGAGAAATACA 501
DB 93 GlyProHisSerAspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThr 112
QY 502 CTGTTTATTCGAAATCCCAACTATCAATAGACAGCAGCTTAAATGCTCTCTGG 561
DB 113 LeuPheTyrSerGluLeuProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrp 132
QY 562 AAGCCTCTTTTGGATCTTTTTCAGGCAACATCGGACTATGGAATGTATTCTCGAGAGAA 621
DB 133 LysProLeuLeuAspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGlu 152
QY 622 GAACATATTAAAGAGAAAGCAATTCGAAACAGTCGGAATGCTTTTACGATTTATCAC 681
DB 153 GluLeuLeuArgGluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHis 172
QY 682 CAAGGAAGTGGAAACATTTCTGTTTCAAGCCGCTAGTGAATTTATCAGTAAAGATGGA 741
DB 173 GlnGlySerGlyThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGly 192
QY 742 GGGCCACAGGATTTACGCAACAACTTTAAGGCCCAATCTAGTGGAAACTAGTGTGCC 801
DB 193 GlyProGlnGlyPheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysPro 212
QY 802 AACATACGATGGATCCAAATATTATGCCCGCTGATCCAGACTGGATGCTTTTATCAT 861
DB 213 AsnIleArgMetAspProLysLeuCysProAlaAspProAspTyrIleAlaPheIleHis 232
QY 862 AGCAACGATATTGATATCTAAACATCGTAAACCAGAGAAAGAGACTCACTTATGTG 921

DB 233 SerAsnAspIleTrpIleSerAsnIleValThrArgGluGluArgArgLeuThrTyrVal 252
QY 922 CACAATGAGCTAGCCACACATCGAAGAAGATCCAGATCAGCTGAGTGCCTACTTTGTT 981
DB 253 HisAsnGluLeuAlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheVal 272
QY 982 CTCCTAAGAAAGAAATTTGATAGATATTCGGCTATTGGTGTGCTCCAAAAGCTGAACAAC 1041
DB 273 LeuGlnGluGluPheAspArgTyrSerGlyTyrTrpCysProLysAlaGluThrThr 292
QY 1042 CCAGTGGTGGTAAATAATCTTAGAATCTTATGAGAAAAATGATGAATCTGAGCTGCAA 1101
DB 293 ProSerGlyGlyLysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGlu 312
QY 1102 ATTATTATGTTTACATCCCTTATGTTGAAAACAAGGAGGAGAGTCAATTCCTGTTATCCT 1161
DB 313 IleIleHisValThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrPro 332
QY 1162 AAAACAGGTACAGCAAAATCCTTAAAGTCACTTTTAAGATGTCAGAAAAATATGATGATCT 1221
DB 333 LysThrGlyThrAlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAla 352
QY 1222 GAAGGAAGCATCATAGATGTCATAGATAAGGAAGTCAATTAACCTTTTGGATTTCTATTT 1281
DB 353 GluGlyArgIleIleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPhe 372
QY 1282 GAAGGAGTTGAATATATTCAGAGCTGGATGGAATCCTCGAGGAAAATATGCTTGGTCC 1341
DB 373 GluGlyValGluTyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSer 392
QY 1342 ATCCTACTAGATGCTCCAGACTCGCTACAGATGATGTTGATCTCAGCTGAATTTATTT 1401
DB 393 IleLeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPhe 412
QY 1402 ATCCAGTAGAAGATGATGTTATGGAAGGAGAGACTCATTGAGTCACTGATGCTTCT 1461
DB 413 IleProValGluAspAspValMetGluArgGlnArgLeuIleGluSerValProAspSer 432
QY 1462 GTGAGCCCATTAATATCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATC 1521
DB 433 ValThrProLeuIleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIle 452
QY 1522 TTTTCATGTTTTTCCCAAGTCAAGAGAGAAATGAGTTATTTTTCCTCTGTAATGC 1581
DB 453 PheHisValPheProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCys 472
QY 1582 AAAACAGGTTTCGTCATTTATACAAAATACATCTATTTTAAAGGAAGCAATATATAA 1641
DB 473 LysThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLys 492
QY 1642 CGATCCAGTGGTGGCTGCTGCTCCCAAGTGAATTCAGTGTCTTATCAAAGAGAGAGATA 1701
DB 493 ArgSerSerGlyGlyLeuProAlaProSerAspPheLysCysProIleLysGluLeu 512
QY 1702 GCATTAACAGTGGTGAAGTTCCTTGGCCGCGCATGATCTTAATATCCAGTTCAT 1761
DB 513 AlaIleThrSerGlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAsp 532
QY 1762 GAAGTCAGAGGCTGGGTATATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACCTG 1821
DB 533 GluValArgArgLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeu 552
QY 1822 TACGTAGTCAGTTACGTAATCTCTGGAGAGGTGACAGGCTCAGCTGACCGTGGCTACTCA 1881
DB 553 TyrValValSerTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSer 572
QY 1882 CATTTCTTGTGCTGATCAGTCAGCTGCTGCTCTTTTATAGTAAGTATAGTAAACAGAG 1941
DB 573 HisSerCysCysIleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLys 592
QY 1942 AATCCACACTGTGTGCTCCCTTTTCAAGCTATCAAGTCTGGAAGATGACCAACTTGCATA 2001
DB 593 AsnProHisCysValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLys 612

CC Zaa is neither Pro nor hydroxyproline.
 CC -|- ENZYME REGULATION: Inhibited by zinc. Inhibited by the serine
 CC proteinase inhibitor 4-(2-aminoethyl)benzenesulphonyl fluoride
 CC (AEBSP), and by di-isopropylfluorophosphate. Specifically inhibited
 CC by isindoline derivatives (By similarity).
 CC -|- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -|- SIMILARITY: Belongs to the peptidase S9B family. DPPIV subfamily.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 CC EMBL; AK016546; BAB30295.2; -; mRNA.
 CC EMBL; BC043124; AAH43124.1; -; mRNA.
 CC EMBL; BC059222; AAH59222.1; -; mRNA.
 CC MEROPS; S09.018; -;
 CC Ensembl; ENSMUSG00000032393; Mus musculus.
 CC MGI; MGI:1921638; Dpp8.
 CC InterPro; IPR001375; Peptidase_S9.
 CC InterPro; IPR002469; Peptidase_S9B.
 CC InterPro; IPR000379; Ser_estr_1.
 CC Pfam; PF00930; DPPIV_N; 1.
 CC Pfam; PF00326; Peptidase_S9; 1.
 CC AminoPeptidase; Hydrolase; Phosphorylation; Protease; Serine protease.
 FT ACT_SITE 749 749 Charge relay system (By similarity).
 FT ACT_SITE 827 827 Charge relay system (By similarity).
 FT ACT_SITE 859 859 Charge relay system (By similarity).
 FT MOD_RES 325 325 Phosphotyrosine.
 FT CONFLICT 87 87 G -> R (in Ref. 1).
 SQ SEQUENCE 892 AA; 102186 MW; 59081CD9792E03ED CRC64;

Alignment Scores:

Pred. No.: 0 Length: 892
 Score: 4559.50 Matches: 849
 Percent Similarity: 97.9% Conservative: 19
 Best Local Similarity: 95.7% Mismatches: 18
 Query Match: 82.1% Indels: 1
 DB: 1 Gaps: 1

US-10-825-632-2 (1-3120) x DPPE_MOUSE (1-892)

Qy 202 GGAATATGCAATGCGAGCAGCAATGGAACAGAACACAGCTGGGTGATATTTGAA 261
 Db 6 GlyArgCysAsnMetAlaAlaMetGluThrGluGlnLeuGlyValGluPheGlu 25
 Qy 262 ACTCGGACTGTGAGNG---AATATTGAATCACAGGATCGGCTTAATTCGAGCCTTT 318
 Db 26 ThrAlaGluCysGluGlyAsnGlyGluSerGlnAspArgProLysLeuGluProPhe 45
 Qy 319 TATGTCGAGCGGTATTCCTGGAGTCAGCTTAAAGAGCTGCTTCCGATACCAGAAATAT 378
 Db 46 TyrValGluArgTyrSerTrpSerGlnLeuLysLeuAlaAspThrArgLysTyr 65
 Qy 379 CATGGCTACATGATGGCTAAGCACCACATGATTTTCATGTTTGTCAAGAGGAATGATCCA 438
 Db 66 HisGlyTyrMetMetAlaLysAlaProHisAspPheMetPheValLysArgThrAspPro 85
 Qy 439 GATGGACCTCATTCAGACAGAACTTATTACCTTGGCATGCTTGGTGAGAACAGAGAAAT 498
 Db 86 AspGlyProHisSerAspArgValTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsn 105
 Qy 499 ACATGTTTTTATCTGAAATTTCCAAACTATCAATATAGACAGCAGTCTTAATGCTCTCT 558
 Db 106 ThrLeuPheTyrSerGluLysProLysThrIleAsnArgAlaAlaValLeuMetLeuSer 125
 Qy 559 TGGAGGCTCTTTTGGATCTTTTTCAGGCAACTGAGCACTATGGAATGATTTCTCGAA 618
 Db 126 TrpLysProLeuLeuAspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGlu 145
 Qy 619 GAAGACTATTAAAGAGAAAGAAACCATTTGGACACATCGGAATGCTTCTTACGATTAT 678
 Db 11eThrIleThrSerGlyGluTrpGluValLeuGlyArgHisGlySerAsnIleTrpVal 525

Db 146 GluGluLeuLeuArgGluArgLysArgIleGlyThrValGlyIleAlaAlaTyrAspTyr 165
 Qy 679 CACCAGAGAGTGGAAATTTCTGTTTCAAGCCGCTAGTGGAAATTTATCAGTAAAGAT 738
 Db 166 HisProGlySerGlyThrPheLeuPheGlnAlaGlySerGlyIleTyrHisIleLysAsp 185
 Qy 739 GGAGGGCCACAGAGATTATCGCAACAACTTTTAAGGCCCAATCTTAGTGGAACTAGTTGT 798
 Db 186 GlyGlyProHisGlyPheThrGlnProLeuArgProAsnLeuValGluThrSerCys 205
 Qy 799 CCCAATACAGATGGATCAAAATATATCCCGCTGATCCAGACTGATGCTTATATA 858
 Db 206 ProAsnIleArgMetAspProLysLeuCysProAlaAspProAspTrpIleAlaPheIle 225
 Qy 859 CATAGCAACATATTTGGATATCTACATCGTAACAGAGAGAGAGAGAGACTCACTTAT 918
 Db 226 HisSerAsnAspIleTrpIleSerAsnLeuValThrArgGluArgIleThrTyr 245
 Qy 919 GTGCACAATCAGCTAGCCCAACATGGAAGAGATGCCAGATCAGCTGGAGTGGCTACCTTT 978
 Db 246 ValHisAsnGluLeuAlaAsnMetGluGluAspProArgSerAlaGlyValAlaThrPhe 265
 Qy 979 GTTCTCCAAGAGAAATTTGATAGATATCTGGCTATTGGTGTGTCCTCAAAAGCTGAACA 1038
 Db 266 ValLeuGlnGluLeuPheAspArgTyrSerGlyTyrTrpTyrCysProGluAlaGluArg 285
 Qy 1039 ACTCCAGTGGTGGTAAAAATTTAGAAATTTCTATATGAAGAAAAATGAAATCTGAGTG 1098
 Db 286 ThrProSerGlyGlyLysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluVal 305
 Qy 1099 GAAATATTATCATGTTACATCCCTATGTTGGAAAACAGAGGGGAGAGATTCATTCGGTTAT 1158
 Db 306 GluIleIleHisValThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyr 325
 Qy 1159 CCTAAACAGGTACAGCAATCTTAAGTCACTTTTAAGATGTCAGAAATATGATTCAT 1218
 Db 326 ProLysThrGlyThrAlaAsnProLysValThrPheLysMetSerGluIleValValAsp 345
 Qy 1219 GCTCAAGAGAGGATCATAGATGTCTAGATAAGGAACATAATCAACCTTTTGGAGATTTCTA 1278
 Db 346 AlaAlaGlyGlyIleIleAspValIleAspLysGluLeuValGlnProPheGluLeu 365
 Qy 1279 TTTCAAGAGTGAATATATTGTCAGAGTGGATGGACTCTCTGAGGAGAAATATGCTTGG 1338
 Db 366 PheGluGlyValGluTyrIleAlaArgAlaGlyTrpThrProGluGlyLysHisAlaTrp 385
 Qy 1339 TCCATCTCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTTCATCTCACTCAATTA 1398
 Db 386 SerIleLeuLeuAspArgSerGlnThrHisLeuGlnIleValIleLeuIleSerProGluLeu 405
 Qy 1399 TTTATCCAGTGAAGATGATGTTATGGAAGAGCAGAGACTCATTCAGTTCAGTGCCTGAT 1458
 Db 406 PheIleProValGluAspAspAlaMetAspArgGlnArgLeuIleGluSerValProAsp 425
 Qy 1459 TCTGTGAGCGCACTAATATCTATGAAGAAACACAGACATCTGTGATAAATATCCATCAC 1518
 Db 426 SerValThrProLeuIleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAsp 445
 Qy 1519 ATCTTTTCATGTTTTCCTCCAAAGTCCAGAGAGAAATGAGTTTATTTTTCCTCTGAA 1578
 Db 446 IlePheHisValPheProGlnThrHisGluAspGluIleGluPheIlePheAlaSerGlu 465
 Qy 1579 TGCAAAACAGGTTTTCCTGCTATTTATACAAATTTACATCTATTATTAAGAGAAAGCAATAT 1638
 Db 466 CysLysThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyr 485
 Qy 1639 AAACGATCCAGTGGTGGCTGCTGCTGCTCCAAAGTATTTCAGTGTCTTATCAAGAGAG 1698
 Db 486 LysArgSerSerGlyGlyLeuProAlaProSerAspPheLysCysProIleLysGlu 505
 Qy 1699 ATAGCAATTTACAGTGGTGAATGGAGTTCTTGGCCGCGCATGGATCTTAATATCAAGTT 1758
 Db 506 IleThrIleThrSerGlyGluTrpGluValLeuGlyArgHisGlySerAsnIleTrpVal 525

[illegible]

NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 272-863 (ISOFORM 2), AND
 PARTIAL NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 5).
 TISSUE=Glial tumor, Ovary, Spleen, and Trachea;
 PubMed=14702039; DOI=10.1038/ng1285;
 Oca T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
 Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
 Sekine M., Obayashi M., Nishi T., Shibehara T., Tanaka T., Ishii S.,
 Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,
 Nagahari K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,
 Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,
 Sugawara M., Takahashi N., Kanda K., Yokoi T., Furuya T., Kikkawa E.,
 Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M.,
 Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,
 Fujimori K., Tanai H., Kimata S., Watanabe M., Hirooka S., Chiba Y.,
 Ishida S., Ono Y., Takiguchi S., Watanabe S., Yoshida M., Hotsuta T.,
 Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,
 Kusano Y., Togiyasu S., Komai F., Hara R., Takeuchi K., Arita M.,
 Imose N., Mutsaers K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
 Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
 Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
 Nakagawa S., Sanoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
 Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
 Yamazaki M., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
 Fujimori Y., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
 Ono T., Yamada K., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
 Kawabata A., Hikiji T., Noguchi S., Itoh T., Shigeta K., Serba T.,
 Oktani R., Kawakami T., Noguchi S., Mizuno T., Morinaga M., Sasaki M.,
 Matsumura K., Nakajima Y., Hatada H., Watanabe M., Komatsu T.,
 Togashi T., Oyama M., Hata H., Watanabe M., Takahashi Y., Nakagawa K.,
 Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Yamashita R.,
 Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
 Nakai K., Yada T., Nakamura Y., Ohara O., Isoigai T., Sugano S.;
 "Complete sequencing and characterization of 21,243 full-length human
 cDNAs.";
 Nat. Genet. 36:40-45(2004).
 [6]
 NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 209-863 (ISOFORM 4), AND
 NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 298-863 (ISOFORM 2).
 TISSUE=Melanoma;
 The German cDNA consortium;
 Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
 -!- FUNCTION: dipeptidyl peptidase that cleaves off N-terminal
 dipeptides from proteins having a Pro or Ala residue at position
 2.
 -!- CATALYTIC ACTIVITY: Release of an N-terminal dipeptide, Xaa-Yaa-[
 Zaa-, from a polypeptide, preferentially when Yaa is Pro, provided
 Zaa is neither Pro nor hydroxyproline.
 -!- ENZYME REGULATION: Inhibited by the serine proteinase inhibitor 4-
 (2-aminoethyl)benzenesulfonyl fluoride (AEBSF), and by di-
 isopropylfluorophosphate.
 -!- BIOPHYSICOCHEMICAL PROPERTIES:
 Kinetic parameters:
 KM=161 uM for Ala-Pro-AMC;
 KM=180 uM for Ala-Pro-AFC;
 pH dependence:
 Optimum pH is 7.5-8.5. Little activity below pH 6.5;
 -!- SUBCELLULAR LOCATION: Cytoplasmic; cytosol.
 -!- ALTERNATIVE PRODUCTS:
 Event=Alternative splicing; Named isoforms=5;
 Name=1; Synonyms=Short;
 IsoId=Q86T12-1; Sequence=Displayed;
 Name=2; Synonyms=Long;
 IsoId=Q86T12-2; Sequence=VSP_013865;
 Note=Incomplete sequence;
 Name=3;
 IsoId=Q86T12-3; Sequence=VSP_013867, VSP_013868;
 Name=4;
 IsoId=Q86T12-4; Sequence=VSP_013869;
 Name=5;
 IsoId=Q86T12-5; Sequence=VSP_013865, VSP_013866;
 Note=Incomplete sequence;
 -!- TISSUE SPECIFICITY: Ubiquitously expressed, with highest levels in
 liver, heart and muscle, and lowest levels in brain.

-!- SIMILARITY: Belongs to the peptidase S9B family. DPPIV subfamily.
-!- CAUTION: Ref.6 (CAD39039) sequence differs from that shown due to framehifts in positions 432 and 460.

This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.

CC	ENBL; AF452102; AAL47179.1; -; mRNA.
CC	ENBL; AY172660; AAO7282.1; -; mRNA.
CC	ENBL; AF542510; AAO73880.2; -; mRNA.
CC	ENBL; AY374518; AAQ83119.1; -; mRNA.
CC	ENBL; BC000970; AAH00970.1; -; mRNA.
CC	ENBL; BC037948; AAB37948.1; -; mRNA.
CC	ENBL; AK054656; BAB70784.1; ALT INIT; mRNA.
CC	ENBL; AK075030; BAC11382.1; -; mRNA.
CC	ENBL; AK131100; BAC85150.1; -; mRNA.
CC	ENBL; AK131499; BAD18643.1; ALT INIT; mRNA.
CC	ENBL; AL834376; BAD39039.3; ALT FRAME; mRNA.
CC	ENBL; CF627380; CAH10477.1; -; mRNA.
CC	MEROPS; S09.019; -;
CC	Ensembl; ENSG00000142002; Homo sapiens.
CC	HGNC; HGNC:18648; DPP9.
CC	MIM; 608258; -.
CC	InterPro; IPR001375; Peptidase_S9.
CC	InterPro; IPR002469; Peptidase_S9B.
CC	InterPro; IPR000379; Ser estsr.
CC	Pfam; PF00930; DPPIV_N; I.
CC	Pfam; PF00326; Peptidase_S9; 1.
KW	Alternative splicing; Aminopeptidase; Hydrolase; Protease;
KW	Serine protease.
FT	ACT_SITE 730 730 Charge relay system (By similarity).
FT	ACT_SITE 808 808 Charge relay system (By similarity).
FT	ACT_SITE 840 840 Charge relay system (By similarity).
FT	VARSPLIC 1 1 M -> LGRVPFCVRRCRCPPLPPLPGSQSRANRREAPLD PCRAQQGRPTKTSVSHACSWNGGSLDDEGTGTPALLRSAE RLMKAVKKLRDKENTGSWRFSLSNSEGAER (in isoform 2 and isoform 5). /FTid=vsp_013865.
FT	VARSPLIC 3 139 Missing (in isoform 5). /FTid=vsp_013866.
FT	VARSPLIC 650 674 QLVNNFSGIKYLRNLTLASLGAV -> SAHLPPPPHH PPEDSPSLKCQL (in isoform 3). /FTid=vsp_013867.
FT	VARSPLIC 675 863 Missing (in isoform 3). /FTid=vsp_013868.
FT	VARSPLIC 832 858 Missing (in isoform 4). /FTid=vsp_013869.
FT	CONFLICT 204 204 I -> N (in Ref. 3).
FT	CONFLICT 571 571 C -> W (in Ref. 5; BAC85150).
FT	CONFLICT 709 709 L -> P (in Ref. 5; BAD18643).
FT	CONFLICT 753 753 G -> C (in Ref. 5; BAB70784).
SQ	SEQUENCE 863 AA; 98263 MW; 40FE0B78E26CDED5 CRC64;

Alignment Scores:

Score:	7.53e-204	Length:	863
Percent Similarity:	2870.00	Matches:	517
Best local Similarity:	77.5%	Conservative:	134
Query Match:	61.5%	Mismatches:	187
Db:	51.7%	Indels:	2
	1	Gaps:	2

US-10-825-632-2 (1-3120) x DPP9_HUMAN (1-863)

QY 316-TTTTATTGTCGGCGGTATTCCTGGAGTCAGCTTAAGAAGCTGCTTGCGCATACAGAAA 375S

Db 24 PheGlnValGlnIysHisSerTrpAspGlyLeuArgSerIleIleHisGlySerArglys 43S

QY 376 TATCATGGCTACATGATGCCTAAGCACACCATGATTTTCATGTTTGTGTGAAGCAAGATGAT 435S

Db 44 TyrSerGlyLeulleValAenIylsAlaproHisAspPheGlnPheValGlnIysThrAsp 63S

QY 436 CCAGATGACCTATTACAGACAGAACTTATTACCTGCCATGCTGGTGAGACAGAGAA 495
DB 64 GluSerGlyProHisSerHisArgLeuTyrTyrLeuGlyMetProTyrGlySerArgGlu 83
QY 496 AATACACGTGTTTATTCGAAATCCCAAAATCTATCAATAGACAGACAGCTCTTAATGCTC 555
DB 84 AnSerLeuLeuLeuSerGluLeuProLysValArgLysGluAlaLeuLeuLeuLeu 103
QY 556 TCTTGAAGCCCTCTTTGGATCTTTTCAGGCAACACTGCACTATGGAATGATTTCTGA 615
DB 104 SerTyrLysGlnMetLeuAspHisPheGlnAlaThrProHisGlyValTyrSerArg 123
QY 616 GAAGAAGAACTATTAGAGAAAGAAACCGCATTTGGAACTGCGAATCTCTTACCAT 675
DB 124 GluGluGluLeuLeuArgGluArgLeuGlyValPheGlyIleThrSerTyrAsp 143
QY 676 TATCAACCAAGGAAGTGGAAACATTTCTCTTTCAAGCCGGTAGTGGAAATTTATCACGTAA 735
DB 144 PheHisSerGluSerGlyLeuPheLeuPheGlnAlaSerAsnSerLeuPheHisCysArg 163
QY 736 GATGAGGGCCACAGGATTTACGCAACAACTTTAAGGCCCAATCTAGTGAAACTAGT 795
DB 164 AspGlyGlyLysAsnGlyPheMetValSerProMetLysProLeuGluIleLysThrGln 183
QY 796 TGTCCCAACATACCGATGGATCCAAAATATTATGCCGCTGATCCAGACTGGATCTTTT 855
DB 184 CysSerGlyProAspMetAspProLysIleCysProAlaAspProAlaPhePheSerPhe 203
QY 856 ATACATAGCAACGATATTGTGATATTAACATCGTAAACAGACAGAGAGAGACTCACT 915
DB 204 IleAsnAsnSerAspLeuTyrValAlaAsnIleGluThrGlyGluLysArgLeuThr 223
QY 916 TATGTCACATAGCTAGCCACATGGAAGAGATGCCAGATCAGCTGAGTCCCTACC 975
DB 224 PheCysHisGlnGlyLeuSerAsnValLeuAspProLysIleGluThrGlyGluLysArgLeuThr 243
QY 976 TTTGTTCTCCAAAGAAATTTGATGATATTCCTGGCTATTCGGCTATTCGGTGTCTCAAAAGCTGAA 1035
DB 244 PheValIleGlnGluLeuPheAspArgPheThrGlyTyrTrpTyrCysProThrAlaSer 263
QY 1036 ACACTCCAGCTGGTGTGTTTAAATTTCTAGAAATTTCTATATGAAGAAATGATGAATCT 1092
DB 264 TrpGluGlySerGluGlyLeuLysThrLeuArgIleLeuTyrGluGluValAspGluSer 283
QY 1093 GAGGTGAAATTTATTCATGTTACATCCCTATGTTGGAAACAAAGGAGGCGAGATTCATTC 1152
DB 284 GluValGluValIleHisValProSerProAlaLeuGluGluLysThrAspSerTyr 303
QY 1153 CGTTATCTCTAAACAGGTACAGCAAAATCTCTAAAGTCATCTTTAAGATGTCAAGAAATATG 1212
DB 304 ArgTyrProArgThrGlySerLysAsnProLysIleAlaLeuLysLeuAlaGluPheGln 323
QY 1213 ATTGATGCTGAAGGAGATCATAGATGTATAGATAAGCAACTTAATTCACCTTTTGAG 1272
DB 324 ThrAspSerGlnGlyLysIleValSerThrGlnGluLysGluLeuValGlnProPheSer 343
QY 1273 ATTCATTATGAAGGAGTTGAATATATTGCCAGAGCTGGATGACTCTCTGAGGGAATAT 1332
DB 344 SerLeuPheProLysValGluTyrIleAlaArgAlaGlyTyrThrArgAspGlyLysTyr 363
QY 1333 GCTTGTGTCCTACTAGATCCCTCCAGACTCCGCTACATAGTACTGATCTCACT 1392
DB 364 AlaTrpAlaMetPheLeuAspArgProGlnGlnTrpLeuGlnLeuValLeuLeuProPro 383
QY 1393 GAATATTTATCCAGTAGAAGATGATGTTATGGAAGGACAGACTCAITGATGCTAGTG 1452
DB 384 AlaLeuPheIleProSerThrGluAsnGluGluGlnArgLeuAlaSerAlaArgAlaVal 403
QY 1453 CCGTATTTCTGTGACGCCACTAATTTATCTATGAAGAAACCAACAGACATCTGGAATAATTC 1512
DB 404 ProArgAsnValGlnProTyrValValTyrGluGluValThrAsnValTrpIleAsnVal 423

QY 1513 CATGACATCTTTCATGTTTTCCTCCAAAGTCAC---GAGAGAGAAATGAGTTTATTTT 1569
DB 424 HisAspIlePheTyrProPheProGlnSerGluGlyGluAspGluLeuCysPheLeuArg 443
QY 1570 GCCTCTGAATGCAAAACAGGTTTCCGTCATTATATACAAAATATCATCTATTTTAAAGGAA 1629
DB 444 AlaAsnGluCysLysThrGlyPheCysHisLeuTyrLysValThrAlaValLeuLysSer 463
QY 1630 AGCAAAATATAACGATCCAGTGGTGGCTGCTCCAAAGTATTTCAGTGTCTCTATC 1689
DB 464 GlnGlyTyrAspTyrSerGluProPheSerProGlyGluAspGluPheLysCysProIle 483
QY 1690 AAAGAGAGATAGCAATATACAGAGTGGTGAATGGGAAGTTCTTGGCCGCACTGATCTAAT 1749
DB 484 LysGluGluIleAlaLeuThrSerGlyGluTyrGluValLeuAlaAlaGlyHisGlySerLys 503
QY 1750 ATCCAAAGTGTGAAGAGGCTGATATATTTTGAAGGACCAAGACTCCCTCTTAA 1809
DB 504 IleTrpValAsnGluLysLeuValTyrPheGlnGlyThrLysAspThrProLeu 523
QY 1810 GAGCATCACTCTGATAGTACGTTACGTAATCTCGAGAGGTGACAGAGCTGACTGAC 1869
DB 524 GluHisHisLeuTyrValValSerTyrGluAlaAlaGlyGluIleValArgLeuThrThr 543
QY 1870 CTTGGCTACTCATCTTCTGTCATCAGTCAGCTGCTGACTTCTTTTAAAGTAACTAT 1929
DB 544 ProGlyPheSerHisSerCysSerMetSerGlnAsnPheAspMetPheValSerHisTyr 563
QY 1930 AGTAACCAAGAAATCCAGCTGCTGCTCCCTTTACAGACTATCACTCTGAGAGTAC 1989
DB 564 SerSerValSerThrProProCysValHisValTyrLysLeuSerGlyProAspAsp 583
QY 1990 CCACTTTGCAAAACAAAGAAATTTTGGCCCACTTTTGGATTTCAGAGGCTCTCTTCT 2049
DB 584 ProLeuHisLysGlnProArgPheTrpAlaSerMetMetGluAlaAlaSerCysProPro 603
QY 2050 GACTATACCTCTCCAGAAATTTCTCTTTTGAAGTACTACTCGAATTTACATTTGATGGG 2109
DB 604 AspTyrValProProGluIlePheHisPheHisThrArgSerAspValArgLeuTyrGly 623
QY 2110 ATGCTCTCAAGCCCTCATGATCTACAGCTGGAAAGAAATATCTACTGTGCTGTTTCATA 2169
DB 624 MetIleTyrLysProHisAlaLeuGlnProGlyLysLysHisProThrValLeuPheVal 643
QY 2170 TATGTTGCTCTCAGGTCAGTTCGTTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGC 2229
DB 644 TyrGlyGlyProGlnValGlnLeuValAsnAsnSerPheLysGlyIleLysTyrLeuArg 663
QY 2230 TTGAATACCTAGCTCTCTAGTTATGTTGTTTGTAGTATAGATAGACAAACAGGCTCTGT 2289
DB 664 LeuAsnThrLeuAlaSerLeuGlyTyrAlaValValIleAspGlyArgGlySerCys 683
QY 2290 CACCAGGCTTAAATTTGAAGGCGCTTTAAATATATAAATGGTCAAAATAGAAATGAC 2349
DB 684 GlnArgGlyLeuArgPheGluGlyAlaLeuLysAsnGlnMetGlyGlnValGluIleGlu 703
QY 2350 GATCAGGTGGAAGGACTCCATATCTAGCTTCTCGATATGATTTTCATGCTTAGATCTGT 2409
DB 704 AspGlnValGluGlyLeuGlnPheValAlaGluLysTyrGlyPheIleAspLeuSerArg 723
QY 2410 GTGGGCACTCAGGCTGCTCTATGAGGATACCTCTCCCTGATGCGATTAATGACAGAG 2469
DB 724 ValAlaIleHisGlyTyrPheSerTyrGlyPheLeuSerLeuMetGlyLeuIleHisLys 743
QY 2470 TCAGATATCTTCAGGTTGCTATTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTTATGAT 2529
DB 744 ProGlnValPheLysValAlaIleAlaGlyAlaProValThrValTrpMetAlaTyrAsp 763
QY 2530 ACAGGATACAGGAACGTTATATGAGGTACCTCCCTGACCAAGTAAAGAGGCTTACTTAA 2589
DB 764 ThrGlyTyrThrGluArgTyrMetAspValProGluAsnAsnGlnHisGlyTyrGluAla 783
QY 2590 GGATCTGTGGCCATGCAAGCAGAGAAAGTTCTCCCTCTGAAACAAATCGTTTACTGTCTTAA 2649

```

Db      784 GlySerValAlaLeuHisValGluLysLeuProAsnGluProAsnArgLeuLeuLeu 803
Qy      2650 CATGGTTCTCGAGCAAGATGCTCCATTTCCACATACCACTAGTATTACTAGCTTTTAA 2709
Db      804 HisGlyPheLeuAspGluAsnValHisPheHisThrAsnPheLeuValSerGlnLeu 823
Qy      2710 GTGAGGCTCGAAAGCCATATGATTACAGATCTATCTCAGAGAGACACACACATAAGA 2769
Db      824 IleArgAlaGlyValProTyrGlnLeuGlnIleTyrProAsnGluArgHisSerIleArg 843
Qy      2770 GTTCTGGAATCGGAGAACATTATGAACTGATCTTTTGGCACTACCTTCAGAAAACCTT 2829
Db      844 CysProGluSerGlyGluHisTyrGluValThrLeuLeuHisPheLeuGlnGluTyrLeu 863

RESULT 4
DPP9_MOUSE
ID      DPP9_MOUSE STANDARD; PRT; 862 AA.
AC      OB8VGA; O6KAM9; OB8WT9;
DT      13-SEP-2005 (Rel. 48, Created)
DT      13-SEP-2005 (Rel. 48, Last sequence update)
DE      Dipeptidyl peptidase 9 (EC 3.4.14.5) (Dipeptidyl peptidase IX) (DP9)
DE      (Dipeptidyl peptidase-like protein 9) (DPLP9).
GN      Name=Dpp9;
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC      Muridae; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 1).
RC      STRAIN=C57BL/6J; TISSUE=liver, and Olfactory bulb;
RX      MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA      Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA      Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA      Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA      Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA      Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA      Blake J.A., Bradt D., Brusci V., Chothia C., Corbani L.E., Cousins S.,
RA      Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA      Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA      Grimmer S., Gustikoff S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA      Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA      Konagaya A., Kurochkin I.V., Lee Y., Lennard B., Lyons P.A.,
RA      Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA      Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA      Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA      Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA      Sandelin A., Schneider C., Sempile C.A., Setou M., Shimada K.,
RA      Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA      Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA      Wilming L.G., Wrynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA      Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA      Hirozane-Kishikawa T., Kono H., Nakamura M., Sakazume T., Sato K.,
RA      Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA      Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA      Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA      Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA      Birney E., Hayashizaki Y.;
RT      "Analysis of the mouse transcriptome based on functional annotation of
RT      60,770 full-length cDNAs."
RT      Nature 420:563-573(2002).
RN      [2]
RP      NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 2).
RC      TISSUE=Thymus;
RX      PubMed=15449545;
RA      Okazaki N., Kikuno R., Ohara R., Inamoto S., Koseki H., Hiraoka S.,
RA      Saga Y., Kitamura H., Nakagawa T., Nagase T., Ohara O., Koga H.;
RT      "Prediction of the coding sequences of mouse homologues of FLJ genes:
RT      the complete nucleotide sequences of 110 mouse FLJ-homologous cDNAs
RT      identified by screening of terminal sequences of cDNA clones randomly
RT      sampled from size-fractionated libraries."

```

```

RL      DNA Res. 11:127-135 (2004).
RN      [3]
RP      NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 1).
RC      STRAIN=FVB/N; TISSUE=Salivary gland;
RX      MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA      Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA      Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA      Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA      Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA      Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA      Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA      Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
RA      Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA      Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA      Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA      Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A.,
RA      Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA      Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA      Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA      Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA      Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalusz D.E.,
RA      Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT      "Generation and initial analysis of more than 15,000 full-length human
RT      and mouse cDNA sequences."
RL      Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC      -|- FUNCTION: Dipeptidyl peptidase that cleaves off N-terminal
CC      dipeptides from proteins having a Pro or Ala residue at position 2
CC      (By similarity).
CC      -|- CATALYTIC ACTIVITY: Release of an N-terminal dipeptide, Xaa-Yaa-|-
CC      Zaa-, from a polypeptide, preferentially when Yaa is Pro, provided
CC      Zaa is neither Pro nor hydroxyproline.
CC      -|- ENZYME REGULATION: Inhibited by the serine proteinase inhibitor 4-
CC      (2-aminoethyl)benzenesulphonyl fluoride (AEBSF), and by di-
CC      isopropylfluorophosphate (By similarity).
CC      -|- SUBCELLULAR LOCATION: Cytosolic (By similarity).
CC      -|- ALTERNATIVE PRODUCTS:
CC      Event=Alternative splicing; Named isoforms=2;
CC      Name=1;
CC      IsoId=Q8BVG4-1; Sequence=Displayed;
CC      Name=2;
CC      IsoId=Q8BVG4-2; Sequence=VSP_013870, VSP_013871, VSP_013872;
CC      -|- SIMILARITY: Belongs to the peptidase S9B family. DPPIV subfamily.
CC      -|- CAUTION: Ref.2 sequence differs from that shown due to a
CC      frameshift in position 745.
CC      -----
CC      This Swiss-Prot entry is copyright. It is produced through a collaboration
CC      between the European Bioinformatics Institute and the EMBL Outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use as long as its content is in no way modified and this statement is not
CC      removed.
CC      -----
CC      EMBL; AK050021; BAC34034.1; -; mRNA.
CC      EMBL; AK078301; BAC37211.1; -; mRNA.
CC      EMBL; AK131178; BAD21428.1; ALT SEQ; mRNA.
CC      EMBL; BC057631; AAH57631.1; -; mRNA.
CC      MEROPE; S09.019; -.
CC      Ensembl; ENSMUSG00000001229; Mus musculus.
CC      MGI; MGI:2443967; Dpp9.
CC      InterPro; IPR001375; Peptidase_S9.
CC      InterPro; IPR002469; Peptidase_S9B.
CC      InterPro; IPR000379; Ser_estrs.
CC      Pfam; PF00930; DPPIV_N; 1.
CC      Pfam; PF00326; Peptidase_S9; 1.
CC      Alternative splicing; Amino peptidase; Hydrolase; Protease;
KW      Serine protease.
KW      ACT_SITE 729 729 Charge relay system (By similarity).
FT      ACT_SITE 807 807 Charge relay system (By similarity).
FT      ACT_SITE 839 839 Charge relay system (By similarity).
FT      VARSPLIC 1 75 Missing (in isoform 2).
FT      VARSPLIC 748 787 /FTId=VSP_013870.
FT      VARSPLIC VALAGAVTVWYDAYDTGYTERVMYVPPENQQGYEAGSVAL
FT      -> PPHEAPSPSLPATTDPKRWASASSMWEAKPGTASEG
FT      QR (in isoform 2).

```

FT VARSPLIC 788 862 /FTid=VSP_013871.
 FT Missing (in isoform 2).
 FT /FTid=VSP_013872.
 FT D -> Y (in Ref. 1; BAC37211).
 FT CONFLICT 369 546 S -> F (in Ref. 2).
 FT CONFLICT 546 777 Q -> K (in Ref. 1; BAC37211).
 FT CONFLICT 777 862 AA; B1D566E824A834E8 CRC64;
 SQ SEQUENCE 862 AA; 98001 MW; 862
 Alignment Scores:
 Pred. No.: 4,18e-201 Length: 862
 Score: 2833.00 Matches: 511
 Percent Similarity: 76.8% Conservatives: 134
 Best Local Similarity: 60.8% Mismatches: 193
 Query Match: 51.0% Indels: 2
 DB: 1 2 Gaps: 2
 US-10-825-632-2 (1-3120) x DPP9_MOUSE (1-862)
 QY 316 TTTTATGTTGAGCGGTATCTCGTCAGTCAGCTTAAAGAGTCGTCCTCCGATACCAAGAAA 375
 DB 23 PheCysValGlnLysHisSerTrpAspGlyLeuArgSerIleIleHisGlySerArgLys 42
 QY 376 TATCATGGCTACATCGCTAAGCCACCAACATGATTTTCATGTTCTGTAAGAGGATGAT 435
 DB 43 SerSerGlyLeuIleValSerLysAlaProHisaspPheGlnPheValGlnLysProAsp 62
 QY 436 CCAGATGACCTTCATTCAGACAGAACTTATACCTCCCATGCTCGTGAGAACAGAA 495
 DB 63 GluSerGlyProHisSerHisArgLeuTyrTrpLeuGlyMetProTyrGlySerArgGlu 82
 QY 496 AATACACTGTTTATTCGAAATCCCAAACTATCAATAGACGACGATCTTAATGCTC 555
 DB 83 AsnSerLeuLeuTyrSerGluIleProLysLysValArgLysGluAlaLeuLeuLeuLeu 102
 QY 556 TCTTGGAAGCCTCTTTTGGATCTTTTTCAGGCAACACTGGACTATGGAATGATTTCTCGA 615
 DB 103 SerTrpLysGlnMetLeuAspHisPheGlnAlaThrProHisGlyValTyrSerArg 122
 QY 616 GAAGAAGAACTATTAAAGAAAGAAAACGCTATCGAACAGTCGCGAATTCGCTTACGAT 675
 DB 123 GluGluGluLeuLeuArgGluArgLysArgLeuGlyValPheGlyIleThrSerTyrAsp 142
 QY 676 TATCAACAAGGAAGTGGAACTATTCGTTTCAACGCGGTAGTGGAAATTTATACGTAA 735
 DB 143 PheHisSerGluSerGlyLeuPheLeuPheGlnAlaSerAsnSerLeuPheHisCysArg 162
 QY 736 GATGGAGGCCACAAAGGATTTACGCAACAACCTTTAAGGCCCAATCTAGTGGAAACTAGT 795
 DB 163 AspGlyGlyLysAsnGlyPheMetValSerProMetLysProLeuGluIleLysThrGln 182
 QY 796 TGTCACACATACGAGTGGATCCAAATATATCCCGCTGATCCAGACTCGGATTCGTTT 855
 DB 183 CysSerGlyProArgMetAspProLysIleCysProAlaAspProAlaPhePheSerPhe 202
 QY 856 ATACATAGCAACGATATTGGATATCTAACTCGTAACCAAGATGCCAGTGGAGTCGCTACC 915
 DB 203 IleAsnAsnSerAspLeuTrpValAlaAsnIleGluThrGlyGluGluValArgLeuThr 222
 QY 916 TATGTCCACAATGAGCTAGCCCAACATGGAAGAAGATGCCAGTGGAGTCGCTACC 975
 DB 223 PheCysHisGlnGlySerAlaGlyValLeuAspAsnProLysSerAlaGlyValAlaThr 242
 QY 976 TTTGTTCTCCAAAGAAGATTTGATAGATATTCGGCTATTGGCTGCTCCAAAGCTGAA 1035
 DB 243 PheValIleGlnGluPheAspArgPheThrGlyCysTrpTrpCysProThrAlaSer 262
 QY 1036 ACAACTCCAGTGGTGGT---AAAAATTCCTTGAATTCCTATATGAAGAAAATGATGATCT 1092
 DB 263 TrpGluGlySerGluGlyLeuLysThrLeuArgIleLeuTyrGluGluValAspGluSer 282
 QY 1093 GAGGTGGAATATTATCATGTTATCATCCCTATGTTGGAACAACAGGAGGCGAGATTCATTC 1152
 DB 283 GluValGluValIleHisValProSerProAlaLeuGluGluArgLysThrAspSerTyr 302

1153 CCGTATCCTTAAACAGGTACAGCAATCCTTAAGTCACTCTTTTAAGATGTGCAGAAATAATG 1212
 DB 303 ArgTyrProArgThrGlySerLysAsnProLysIleAlaLeuLysLeuAlaGluLeuGln 322
 QY 1213 ATTGATGTGAAGGAGGATCATAGATGTCTATAGTAAGGAACCTTAATTCACCTTTTGAG 1272
 DB 323 ThrAspHisGlnGlyLysIleValSerSerCysGlnLysGluLeuValGlnProPheSer 342
 QY 1273 ATTCTATTGAAGGAGTTGAATATATTCGCCAGACTGGATGGATCTCTGAGGGAATAT 1332
 DB 343 SerLeuPheProLysValGluTyrIleAlaAlaGluArgLysTrpThrArgAspGlyLysTyr 362
 QY 1333 GCTTGGTCCATCCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGTGATCTACCT 1392
 DB 363 AlaTrpAlaMetPheLeuAspArgProGlnGlnArgLeuGlnLeuValLeuLeuProPro 382
 QY 1393 GAATATTATCCCTAGTAGAAGATGATGTTATGAAAGCGAGAGACTCATTTGATGACTAGT 1452
 DB 383 AlaLeuPheIleProAlaValGluSerGluAlaGlnArgGlnAlaAlaAlaArgAlaVal 402
 QY 1453 CCTGATCTGTGACGCCCACTAATATCTATGAGAAAACAACAGACATCTCGATAAATATC 1512
 DB 403 ProLysAsnValGlnProPheValIleTyrGluGluValThrAsnValTrpIleAsnVal 422
 QY 1513 CATGACATCTTTCATGTTTTCCTCCCAAGTCAC--GAAGAGGAAATGAGTTTATTTT 1569
 DB 423 HisAspIlePheHisProPheProGlnAlaGluGlnGlnAspPheCysPheLeuArg 442
 QY 1570 GCCTCTGAATGCAAAACAGTTTCCTGCTATTTATACAAAATTTACATCTATTTTAAAGCAA 1629
 DB 443 AlaAsnGluCysLysThrGlyPheCysHisLeuTyrArgValThrValGluLeuLysThr 462
 QY 1630 AGCAATATTAACGATCCAGTGGCTGGCTGCTCCAGTGATTTCAAGTGCCTATC 1689
 DB 463 LysAspTyrAspTrpThrGluProLeuSerProThrGluAspGluPheLysCysProIle 482
 QY 1690 AAAGAGGAGATAGCAATACAGTGGTGAATGGAGTTCTTCCCGGCGATCGATTAAT 1749
 DB 483 LysGluGluValAlaLeuThrSerGlyGluTrpGluValLeuSerArgHisGlySerLys 502
 QY 1750 ATCAAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGCCACCAAGACTCCCTTTA 1809
 DB 503 IleTrpValAsnGluGlnThrLysLeuValTyrPheGlnGlyThrLysAspThrProLeu 522
 QY 1810 GAGCATCAGCTGTAGCTAGTACGTAAATCTCGAGAGGTGACAGGCTGACTGAC 1869
 DB 523 GluHisHisLeuTyrValValSerTyrGluSerAlaGlyGluIleValArgLeuThrThr 542
 QY 1870 CGTGGCTACTACATCTTTCGTCATCAGTCAGCACTGTGACTCTTCTTATAGTAAGTAT 1929
 DB 543 LeuGlyPheSerHisSerCysSerMetSerGlnSerPheAspMetPheValSerHisTyr 562
 QY 1930 AGTAACCCAGAGAATCCACACTGTGTGTCCTTTTCAAGCTATCAAGTCTCTGAAAGTAC 1989
 DB 563 SerSerValSerThrProCysValHisValTyrLysLeuSerGlyProAspAspAep 582
 QY 1990 CCACTTGCAAAACAAAGGAATTTTGGGCCACCATTTTGGATTTCAGCAGGTCTCTTCTCT 2049
 DB 583 ProLeuHisLysGlnProArgPheTrpAlaSerMetMetGluAlaAlaAsnCysProPro 602
 QY 2050 GACTATCTCTCCAGAAATTTCTTTTGAAGCTACTACTGATTTTACATTTGATGCGG 2109
 DB 603 AspTyrValProGluIlePheHisPheHisThrArgAlaAspValGlnLeuTyrGly 622
 QY 2110 ATGCTCTACAGCTCATGATCTACAGCTCGAAGAAATATCTTACTGTCTGTTCATA 2169
 DB 623 MetIleTyrLysProHisThrLeuGlnProGlyArgLysHisProThrValLeuPheVal 642
 QY 2170 TATGCTGCTCCTCAGTGCAGTGTGGGAATATCGTTTAAAGGATCAAGTATTCGCGC 2229
 DB 643 TyrGlyGlyProGlnValGlnLeuValAsnAsnSerPheLysGlyIleLysTyrLeuArg 662

```
QY 2230 TTGAATACCTAGCCTCTCTAGGTTATGTGGTTATGTAGTGATAGACACAGGGGATCCTCT 2289
DB |||||ThrLeuAlaSerLeuGlyTyrAlaValValValLeuLeuGlySerCys 682
QY 2290 CACCGAGGCTTAAATTTGAAGCGCCTTTAAATATAAATGGTCAATAGAAATGAC 2349
DB |||||GlnArgGlyLeuHisPheGluGlyAlaLeuLysAsnGlnMetGlyGlnValGluLeuGlu 702
QY 2350 GATCAGGTGGAGGACTCCAAATATCTAGCTTCTCGATATGATTTTCAATGCTTAGATCT 2409
DB |||||GlnValValGluGlyLeuGlnTyrValAlaGluLysTyrGlyPheLeuAspLeuSerArg 722
QY 2410 GTGGGCATCCAGCGCTGCTTATCGAGGATACCTCTCCCTGATGGCATTAAATGACAGAG 2469
DB |||||ValAlaLeuHisGlyTyrPheTyrGlyGlyPheLeuSerLeuMetGlyLeuLeuHisLys 742
QY 2470 TCAGATATCTCAGGCTTCTATTCCTGGGCGCCAGTCACTCTGTGGATCTTCTATGAT 2529
DB |||||ProGlnValPheLysValAlaLeuAlaGlyAlaProValThrValTrpMetAlaTyrAsp 762
QY 2530 ACAGGATACACGGAAGCTTATATGGGTCACTCCCTGACACAGATGAACAGGCTTACTTAA 2589
DB |||||ThrGlyTyrThrGluArgTyrMetAspValProGluAsnAsnGlnGlyTyrGluAla 782
QY 2590 GGATCTGTGGCCATGCAAGACAGAAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTA 2649
DB |||||GlySerValAlaLeuHisValGluLysLeuProAsnGluProAsnArgLeuLeuLeu 802
QY 2650 CATGGTTTCTGTGATGAGATGTCATTTTGGCATACCAGTATATTTACTGAGTCTTTTAA 2709
DB |||||HisGlyPheLeuAspGluAsnValHisPhePheHisThrAsnPheLeuValSerGlnLeu 822
QY 2710 GTGAGGCTCGAAGCCATATGATTTACAGATCTATCTCTCAGGAGAGACACAGCAAGA 2769
DB |||||IleArgAlaGlyLysProTyrGlnLeuGlnIleTyrProAsnGluArgHisSerIleArg 842
QY 2770 GTTCTCCTGAATCGGAGAACATTTGAACTGCACTCTTTTGGCACTACCTTCAAGAAAACCT 2829
DB |||||CysArgGluSerGlyGluHisTyrGluValThrLeuLeuHisPheLeuGlnGluHisLeu 862

RESULT 5
Q4SBM6_TETNG PRELIMINARY; PRT; 923 AA.
AC Q4SBM6;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Chromosome 15 SCAF14667, whole genome shotgun sequence.
DE (Fragment).
GN ORFNames=GSTENG0020903001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
CX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Jallion O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anhouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA Biemont C., Skalli Z., Cattolico L., Poulain J., De Bernardinis V.,
RA Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
RA Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Wolff J.N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Winkler P., Lander E.S., Weissbach J., Roest Crolius H.;
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
```

```
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAAE01014667; CAG01956.1; -; Genomic_DNA.
FT NON TER 1 1
FT NON TER 923 923
SQ SEQUENCE 923 AA; 105211 MW; 1397023B2004D009 CRC64;

Alignment Scores:
Pred. No.: 7,75e-201 Length: 923
Score: 2829.50 Matches: 524
Percent Similarity: 71.6% Conservative: 138
Best Local Similarity: 56.6% Mismatches: 190
Query Match: 51.0% Indels: 73
DB: 2 Gaps: 7

US-10-825-632-2 (1-3120) x Q4SBM6_TETNG (1-923)
QY 268 GACTGTGAGGAGAAATTTGAATCACAGGATCGGCCTAAATTTGGAGCCTTTTATGTTGAG 327
DB |||||1 AspSerThrGluValValGluMetGluAspValPro-----SerGlnPheValGln 18
QY 328 CGGTATTCTTGGAGTCAGCTTAAAGCTGTGTCGCGATACACAGAAATATCATGGCTAC 387
DB |||||19 LysHisSerTrpGluGlyLeuArgAspIleIleHisCysSerArgLysAsnSerGlyLe 38
QY 388 ATGATGGCTAAGCACCACATGATTTTCATGTTTGTGAGAGGATGATCCAGATGGACCT 447
DB |||||39 IleAlaAsnLysAlaProHisAspPheGlnPheValGlnLysLysAsnGluAsnGlyPro 58
QY 448 CATTACAGACAGATCTTATTACCTTGGCCTGTGCTGGTGGAGACAGAGAAATACACTGTT 507
DB |||||59 HisSerHisArgLeuTyrTyrLeuGlyMetProTyrGlySerArgLysSerLeuLeu 78
QY 508 TATTCTGAAATTCCTCAATATAGACAGCAGCTCTTAATCTCTCTTGGAGAGCCT 567
DB |||||79 TyrSerGluLeuProLysLysValArgLysGluAlaLeuLeuValLeuSerTrpLysGln 98
QY 568 CTTTGTGATCTTTTTCAGGCAACACCTGCACTATGGAATGTATTCTCGAAGAGAAAGAACTA 627
DB |||||99 MetLeuAspHisPheGlnAlaThrProHisGlnGlyAlaTyrSerArgGluGluGluLeu 118
QY 628 TTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 687
DB |||||119 LeuArgGluArgLysArgLeuGlyAlaPheGlyIleThrSerTyrAspTyrHisAlaGln 138
QY 688 AGTGGACATCTTCTGTTTCAAGCCGCTAGTGAATTTATCAGTAAAAGATGGAGGCGCA 747
DB |||||139 ThrGlyLeuPheLeuPheGlnAlaSerAsnSerLeuPheTyrCysGlnAspGlyGlyGln 158
QY 748 CAAGGATTTACGCAACAA---CCTTTAAGGCCCAATCTAGTGGAAACTAGTTGTCCTCCCA 804
DB |||||159 AsnSerPheIleGlnSerAlaProValLysProValGluIleLysThrGlnCysSerGly 178
QY 805 ATACGGATGATCCAAATTTATCCCGCTGATCCAGACTGGATGCTTTTATCATATAGC 864
DB |||||179 ThrArgMetAspProLysIleCysProAlaAsnProAspPheIleAlaPheIleAsnAsn 198
QY 865 AACCATATTTGGATATCTAACATCTGTAACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 924
DB |||||199 AsnAspLeuTrpValAlaAsnIleLysThrGlyGluArgArgLeuThrPheCysHis 218
QY 925 AATGAGCTAGCCAACTAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 984
DB |||||219 LysGlyAlaAspSerValLysGluAspProLysSerAlaGlyValAlaThrPheValIle 238
QY 985 CAAGAAGATTTGATAGATATTTCTGGCTATTGCTGTGTCCTCAAAAGCTGAAACAACTCCC 1044
DB |||||239 GlnGluGluPheAspArgPheThrGlyTyrTrpTrpSerProSerAlaValGluAspPro 258
```


QY	1045	AGTGTGTTAAATTTCTAGAAATTTCTATATGAAGAAAATGATGAATCTGAGGTGAAATTT	1104
Db	259	AepGlyGlyLysArgValTyrLeuLeuTyrGluGluValAspGluThrGluValGluIle	278
QY	1105	ATTCAATGTTACATCCCTATGTTGGAAAACAAGGAGGCGAGATTCATTCCGTTATCCTAAA	1164
Db	279	IleHieValProSerProAlaLeuGluGluArgLysAlaAspAlaTyrArgTyrProArg	298
QY	1165	ACAGGTACAGCAATCTCTAAAGTCACTTTTAAAGATTCAGAAATATGATGATCCTGAA	1224
Db	299	ThrGlySerLysAsnProGlnAlaThrIleLysLeuValGluIleLysThrAspGlnGln	318
QY	1225	GGAAGG-----	1230
Db	319	GlyArgValSerLysCysArgLeuLeuCysValPheLeuLeuArgSerAspArgAla	338
QY	1231	-----ATCATAGATGTCATAGATAAGCAACTAATTCACACT	1266
Db	339	ValLeuThrCysLeuTyrLeuGlnIleValSerThrGlnAspLysGluLeuAlaValPro	358
QY	1267	TTTGAGATCTATTGAAGAGTTGAAATATATGTCAGAGCTGGATGGACTCTCAGGGA	1326
Db	359	PheThrSerLeuPheProGlyThrGluTyrIleAlaArgValGlyTrpThrSerAspGly	378
QY	1327	AAATAGCTGTGTCATCTCTAGATCGCTCCAGACTCGCTACAGATAGTGTGATC	1386
Db	379	LysTyrGlyTrpAlaAlaLeuLeuAspArgSerGlnArgLysLeuGlnValLeuLeu	398
QY	1387	TCACCTGAATATTATCCAGTAGAAGATGATGTTATGGAAGGAGAGACTCATGAG	1446
Db	399	ProProAlaPhePheValProValThrAspAspProAlaArgArgGlnSerLeuGlu	418
QY	1447	TCAGTGCCTGATTCGTGACCCCACTAATATCTATGAAGAAAACAAGACATCTGGATA	1506
Db	419	AlaValProAspArgThrGlnProTyrValIleTyrGluGluThrThrAspValTrpIle	438
QY	1507	ATATCCATGACATCTTTCATGTTTTCCTCCCAAGTCAAGAGGAGAAATTTGAGTTATT	1566
Db	439	AsnValHisAspIlePheTyrProPheValGlnThrAlaGluAspGluPheThrPheIle	458
QY	1567	TTTGCTCTGAAATGCAAAACAGTTTCGTCTATTATACAAAATTCATCTATTTTAAAG	1626
Db	459	TrpValAsnGluSerLysThrGlyPheSerHisLeuTyrLysIleThrSerValLeuHis	478
QY	1627	-----GAAAGCAAAATATAACGATCCAGTGTGTGG-----	1656
Db	479	ProGlyPheHisCysTrpAlaGluAlaTyrHisIleThrGluGlyAspProGlnArgArg	498
QY	1657	CTGCTGCT-----CCAAAGTGATTTCAAGTGTCTCTATCAAGAGGAG	1698
Db	499	IleProAlaValSerThrAspValProProGlyAspPheLysCysAlaValLysGluGlu	518
QY	1699	ATAGCAATTACCGTGGTGAATGGAGTTCTTGGCGGATGATCTATATATCCAGTT	1758
Db	519	IleThrLeuThrSerGlyGluTrpGluValLeuAlaArgHisGlySerLysIleTrpVal	538
QY	1759	GATGAAGTCAGAGCGTGTATATTTGAGGACCAAGACTCCCTTTTAGAGCATCAC	1818
Db	539	AsnGluSerSerLysLeuValTyrPheGlnGlyThrArgAspThrProLeuGluHisHis	558
QY	1819	CTGTAGTGTAGTTCAGTTACATAATCTCTGAGAGGTGACAGGCTGACTGACCGTGCCTAC	1878
Db	559	LeuTyrValValSerTyrSerProGlyAspValValArgLeuThrLysProGlyPhe	578
QY	1879	TCACATTTCTGCTGATCATGT-----	1899
Db	579	SerHisSerCysSerValSerGlnValLysLysSerLeuGlnSerAspTyrPhePheAsn	598
QY	1900	-----CAGCACTGTGACTTC	1914
Db	599	TyrSerSerIleThrLeuProLeuSerLeuSerSerPheIleTrpGlnAsnPheAspPhe	618
QY	1915	TTTATAAGTAAGTATAGTAACCAAGAAATCCACACTGTGTGTCTCTTACAGCTATCA	1974
Db	619	PheValSerHisTyrSerSerValCysThrProCysValHisValTyrLysLeuAsn	638
QY	1975	AGTCTGGAAGATGACCCCACTTGCAAAACAAGGAATTTTGGGCCACCACTTTTGGATTCA	2034
Db	639	SerSerGluSerAspProLeuHisIleValProGluPheTrpAlaSerMetMetGluSer	658
QY	2035	GCAGTCTCTTCTGACTATCTCTCCAGAAATTTTCTTTTCAAGACTACTACTGGA	2094
Db	659	SerGlyCysProGlyAspTyrSerProGluIlePheAspPheGlnGlyLysSerGly	678
QY	2095	TTTACATGTTATGGGATCTCTCAAGCTCATCATCTACAGCTCGGAAGAATATCTCT	2154
Db	679	PheGlnLeuTyrGlyMetValTyrLysProHisSerLeuGlnProGlyArgLysHisPro	698
QY	2155	ACTGTGCTGTTTCATATATGTTGGTCTCAGGTGCAGTGTGTGTAATATCGTTTAAAGGA	2214
Db	699	ThrValLeuPheValTyrGlyGlyProGlnValGlnLeuValAsnAsnSerPheLysGly	718
QY	2215	GTCAAGTATTTCCGCTTGAATACCTAGCTCTCTAGGTTATGTGGTTGTAGTAGTAGAC	2274
Db	719	MetLysTyrLeuArgLeuAsnThrLeuAlaSerLeuGlyTyrAlaValValIleAsp	738
QY	2275	AACAGGGATCTGTCAACCGAGGGCTTAAATTTGAAGCGCCTTTAAATATAAAATGGGT	2334
Db	739	GlyArgGlySerCysGlnArgGlyLeuGluPheGluSerAlaLeuLysAsnLysMetGly	758
QY	2335	CAAAATAGAATTCACGATCAGGTGAGGACTCCAATATCTAGCTTCTCGATATGATTTTC	2394
Db	759	GlnValGluIleGluAspGlnValGluGlyLeuGlnTyrValAlaGluLysPheAsnPhe	778
QY	2395	ATTGACTTAGATCGTGTGGGATCCACGGCTGGTCTCTATGGAGGATACCTCTCCCTGATG	2454
Db	779	ValAspPheSerArgValAlaIleHisGlyTrpSerTyrGlyGlyPheLeuSerLeuMet	798
QY	2455	GCATTAATCCAGGTCAGATATCTCAGGTCCTATTTGCTGGGGCCCGACCTCCTCTG	2514
Db	799	GlyLeuIleGlnArgProAsnValPheLysLeuAlaIleAlaGlyAlaProValThrVal	818
QY	2515	TGATCTTCTATGATACAGGATACAGCAAGCTTATATGGTCACTCCCTGACCAAGATGAA	2574
Db	819	TrpMetAlaTyrAspThrGlyTyrThrGluArgTyrMetAspValProGluAsnAsnGln	838
QY	2575	CAGGGCTATTACTTAGGATCTGTGCCATGCAAGCAGCAAAAGTTCCCTCTGACCAAAAT	2634
Db	839	GlnGlyTyrGluGluGlySerValAlaLeuHisValAspLysLeuProSerGluProAsn	858
QY	2635	CGTTTACTCTTATACATGTTTCTCTGATGAGAAATGTCATTTTGACATACACAGTATA	2694
Db	859	ArgLeuLeuIleLeuHisGlyPheLeuAspGluAsnValHisPheHisThrAsnPhe	878
QY	2695	TTACTGAGTTTTTTAGTGGGCTGGAAGCCCATATGATTACAGATCTATCTCTCAGGAG	2754
Db	879	LeuValSerGlnIleIleArgAlaGlyLysProTyrGlnLeuGlnValTyrProAsnGlu	898
QY	2755	AGACACAGCATAGAAGTCTCTGAATCGGAGAACATTTATGAATGCATCTTTTGCACCTAC	2814
Db	899	ArgHisSerIleArgCysProGluSerGlyGluHisTyrGluIleMetLeuLeuHisPhe	918
QY	2815	CTTCAAGAAAACCTT	2829
Db	919	LeuGlnGlnTyrLeu	923
RESULT 6			
ID	Q6GR22_XENLA	PRELIMINARY;	PRT; 847 AA.
AC	Q6GR22_XENLA	27, Created)	
DT	05-JUL-2004	(REMBUREL. 27, Last sequence update)	
DE	05-JUL-2004	(REMBUREL. 27, Last annotation update)	
DE	MGC81313	protein.	
GN	Name=MGC81313;		

OS *Xenopus laevis* (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
 OC Xenopodinae; Xenopus; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Embryo;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.P., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Embryo;
 RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
 RA Richardson P.;
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
 RT initiative.";
 RL Dev. Dyn. 225:384-391 (2002).
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Embryo;
 RA Klein S., Gerhard D.S.;
 RL Submitted (May-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC071112; AAH71112.1; -; mRNA.
 DR GO; GO:0003824; F: catalytic activity; IEA.
 DR GO; GO:0008236; F: serine-type peptidase activity; IEA.
 DR GO; GO:0006508; P: proteolysis and peptidolysis; IEA.
 DR InterPro; IPR001375; Peptidase_S9.
 DR InterPro; IPR002469; Peptidase_S9B.
 DR InterPro; IPR00379; Ser. esters.
 DR Pfam; PF00930; DPPIV_N; I.
 DR Pfam; PF00326; Peptidase_S9; 1.
 SQ SEQUENCE 847 AA; 97484 MW; F515F2609B267BAF CRC64;

Alignment Scores:
 Pred. No.: 8.51e-192 Length: 847
 Score: 2707.50 Matches: 495
 Percent Similarity: 74.3% Conservative: 140
 Best Local Similarity: 57.9% Mismatches: 209
 Query Match: 48.8% Indels: 11
 DB: 2 Gaps: 3

US-10-825-632-2 (1-3120) x Q6GR22_XENLA (1-847)

QY 268 GACTGTGAGGAGATATTGATACACAGATCGCGCTAAATTGGAGCCTTTTATGTTGAG 327
 Db 3 AspaLaGluGlyArgValAlaSerHis-----PhelysValGlu 15
 QY 328 CGGTATTCCTGGAGTCAGCTTAAAGCTGTTCCGCGATACCAAGAAATATCATGGCTAC 387
 Db 16 ArgArgSerTrpGluGlyLeuArgAspIleHisGlySerArgLysTrpSerGlyIle 35
 QY 388 ATGATGGCTAAGGACCACCATGATTTTCATGTTTGTGAAGGAGATGATCCAGATGGACCT 447

36 MetValAsnLysAlaProHisAspTyrHisPheGlnSerArgLysGluAspSerGlyThr 55
 448 CATTACAGACAGAAATCTATTACCTTGCATGCTCTGGTGAGAACAGAGAAAATACACTGTTT 507
 56 HisSerHisArgLeuTyrPheLeuGlyMetProTyrGlyThrArgGluAsnThrLeuLeu 75
 508 TATTCTGAAATTCCTCAAACTATCAATAGACAGCAGCTTTAATGCTCTCTTTGGAGCCT 567
 76 TyrSerGluLeuProArgGlnValArgLysGluThrThrLeuLeuLeuSerTrpLysGln 95
 568 CTTTGTGATCTTTTTCAGGCAACACCTGGACTATGAATGATTTCTTCGAGAGAGAACTA 627
 96 LeuLeuGluAsnPheGlnAlaThrProHisHisGlyValTyrSerArgGluGluLeu 115
 628 TTAAGAGAAAAGAACCATTTGAAACAGTCGGAATTTCTTACGATATATCACCAGGA 687
 116 LeuArgGluArgLysArgLeuGlyGlyPheGlyIleThrSerTyrAspPheHisSerLys 135
 688 AGTGGACATTTCTGTTTCAAGCGGTAGTGGAAATTTATCAGTAAAGATGGAGGGCCA 747
 136 ThrGlyPhePheLeuPheProAlaAsnAsnSerLeuTyrTyrCysArgAspGlyGlySer 155
 748 CAAGATTTACGCAACAACTTTAAG3CCCAATCTAGTGGAACTAGTTGTCCCAACATA 807
 156 IleSerPheSerAlaSerProValLysProIleGluIleLysThrLeuCysAspGlyPro 175
 808 CGGATGGATCCAAATTTATGCCCGCTGATCCAGACTGGATGCTTTTATATACATAGCAAC 867
 176 ArgMetAspAlaLysIleCysProAlaAsnProAspPheIleSerPheIleAsnAsnAsn 195
 868 GATATTTGGATATCTAACTCGTAACACGAGAGAAAGGAGACTCACTTATGTGCACAA 927
 196 AspLeuTrpIleGluAsnIleHisThrMetGluGluLysArgLeuThrTyrCysHisLys 215
 928 GAGCTAGCCAACTGGAAGAGATGCCAGATCAGCTGGAGTGCCTACCTTTGTTCTCCAA 987
 216 GlyLeuHisThrValMetAspAspProCysSerAlaGlyValAlaThrPheValIleGln 235
 988 GAAGAATTTGATAGATATCTGCTATTTGGTGTGTCTCCAAAGCTGAAACAACACTCCCAGT 1047
 236 GluGluPheAspArgPheThrGlyTyrTrpTrpSerProThrValValAspLeuProGly 255
 1048 GGTGTGTAATTTCTAGAAATTTCTATGAGAAATGATGAATCTGAGTGGAAATATT 1107
 256 GlyPheSerValLeuArgIleLeuTyrGluGluValAspGluSerAspValGluIle 275
 1108 CATGTTACATCCCTATGTTGAAACAAAGGAGGCGGAGTTCATTCGTTATCTCTAAACA 1167
 276 HisValProSerProAlaLeuGluGluArgLysThrAspAlaTyrArgTyrProArgThr 295
 1168 GGTACAGCAAACTCTAAAGTCACTTTTAAGATGTCGAAATAATGATGATGCTCTGAAGA 1227
 296 GlySerLysAsnProLysIleThrLeuLysLeuSerGluIleThrThrAsnAsnAspGly 315
 1228 AGCATCATAGATGTCATAGATAAGGAACCTAAATTCACCTTTTTCAGATTTCTATTGAAGA 1287
 316 GluIleValAlaAlaGlnAspLysGluLeuValLeuProPheHisSerLeuPheProSer 335
 1288 GTTGAATATATTGCCAGAGCTGGATGGACTCTCGAGGAGAAATATGCTGGTCCATCTTA 1347
 336 AlaGluTyrIleValArgAlaGlyTrpSerGlnAspGlyLysPheAlaTrpAlaMetLeu 355
 1348 CTAGATCCTCCAGACTCGCTACAGATAGTGTGATCTCACCTACCTGAAATATTATCCCA 1407
 356 LeuAspArgProGlnCysLeuLeuGlnValLeuLeuProProGluLeuPheIlePro 375
 1408 GTAGAGATGATGTTATGGAAAGCAGAGACTCATTTGAGTCAGTGCCTGATTTCTGTGAG 1467
 376 ValCysGluAspGluGluGlnArgIleGlnTyrValLysLysValProGlnHisIleGln 395
 1468 CCACATAATTCTATGAGAAACCAACAGACATCTCGGATAAATATCCATCAGATCTTTTCA 1527

D	396	ProTyrValIleTyGluGluLysSerAspValTrpValAsnValHisAspValValHis	415
Q	1528	GTITTTTCCCAAGTCAAGAGAAATGAGTTTATTTTCCCTCTGAATGCAAAACA	1587
D	416	LeuLeuArgGlnGluLysGluAsnGluIleArgPheIleCysSerIleGluSerGlnThr	435
Q	1588	GGTTTCCGTCATTTATACAAATACATCTATTATTAAGAGAAAGCAATATATAACGATCC	1647
D	436	GluPheCysHisLeuTyLeuMetThrSerLeuLeuGlnAsnGluAsnArgGluTrpSer	455
Q	1648	AGTGTGGGCTGCTCTCCAAAGTGAAT---TTCAAGTGCTCTATCAAGAGGAGATAGCA	1704
D	456	GlnMetThrIleProCysThrValAspMetPheLysCysGlnIleAsnGlnGluValAla	475
Q	1705	ATTACAGTGTGAATGGGAAGTCTTCGGCGGATGGATCTAATATCAAGTTGATGAA	1764
D	476	LeuThrSerGlyGluTrpGluValLeuCysArgHisGlyAlaLysValTrpValAsnGlu	495
Q	1765	GTCAAGGCTGGTATATTTTGAAGGACCAAGACTCCCTTTAGAGCATCACCTGTAC	1824
D	496	AlaMetGlnLeuValTyPheGlnGlyThrLysAspThrProLeuGluHisLeuTy	515
Q	1825	GTAGTCAGTTACGTAATCTCGAGAGGTGACAAAGGCTGACTACCGTGGCTACTCACAT	1884
D	516	ValThrSerTyGlnHisProGlyGluValValArgLeuThrGluArgGlyTySerHis	535
Q	1885	TCTTGTCGATCAGTCAGCTGTGCTTCTTTATAGTAAGTATAGTAACCAAGAGAT	1944
D	536	SerCysThrMetSerProAsnPheAspMetPheValSerGlnTySerSerValSerCys	555
Q	1945	CCACACTGTGTGCTCTTTTACAAAGCTATCAAGTCTCGAAGTACCAACCTTGCAAAACA	2004
D	556	ProProCysValHisLeuTyArgLeuAsnGly-----AsnProLeuTyGlnHis	572
Q	2005	AAGGAATTTTGGGCCACCATTTTGGATTCAGCAGGTCTCTTCTGACTATATCTCTCA	2064
D	573	ProGlnPheTrpAlaSerValMetGluAlaLysArgCysProProTyTyValProPro	592
Q	2065	GAATTTTCTCTTTGAAGTACTACTGGAATTTACATTTGATGGATGCTCTCAAGCT	2124
D	593	GluIlePheHisPheGlnAlaSerSerGlnValLysLeuTyGlyMetValTyLysPro	612
Q	2125	CATGATCTACAGCTGGAAGAATAATCTCTACTGCTGCTTTCATATATGGTGTCTCTCAG	2184
D	613	HisAsnLeuValProGlyThrLysHisProThrValLeuPheValTyGlyProGln	632
Q	2185	GTGCAAGTTGGTGAATAATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCTAGCC	2244
D	633	ValGlnLeuValAsnAsnSerPheLysGlyMetLysTyLeuArgLeuAsnThrLeuAla	652
Q	2245	TCTCTAGGTTATGTGTTGTAGTATAGACAAACAGGGGATCTCTGACCGAGGCTTAAA	2304
D	653	HisLeuGlyTyAlaValValIleAspGlyArgGlySerCysHisArgGlyLeuAla	672
Q	2305	TTTGAAGCGCCTTTAAATATAAATGGGTCAAATAGAAATTTGACGATCAGGTGGAAGGA	2364
D	673	PheGluGlyAlaLeuLysAsnLysMetGlyGlnValGluIleGlnAspGlnValGluGly	692
Q	2365	CTCCAATATCTAGCTTCTCGATATGATTTTCATTGATCTTAGATCGTGGGCATCCAGCC	2424
D	693	LeuHisPheValAlaGluArgPheGlyPheValAspLeuAspArgValSerIleHisGly	712
Q	2425	TGTCCTATGAGGATACCTCTCCCTGATGGCATTAAATCAGAGTCAAGATCTTTCAGG	2484
D	713	TrpSerTyGlyGlyPheLeuSerLeuMetGlyIleIleGlnArgProGluValPheLys	732
Q	2485	GTTGCTATTCGTGGGCCCCAGTCACCTCTGTGGATCTTCTATGATACAGGATACCGGA	2544
D	733	ValAlaIleAlaGlyAlaProValThrLeuTrpMetAlaTyAspThrGlyTyThrGlu	752
Q	2545	CGTTATATGGGTCCACCTGACCAAGTAACAGGGCTATTACTTTAGGATCTGTGCCATG	2604
D	753	ArgTyMetGluThrProGluThrAsnGlnMetGlyTyGluAlaGlySerAlaAlaLeu	772

34 AspArgProGlnGlnTrpLeuValLeuValLeuProAlaLeuPheIleProSer 153
1411 GAGATGATGTTGGAAGCGAGAGCTCATTGAGTCCGCTGATCTCTGACGCCA 1470
54 ThrGluAsnGluGluGlnArgLeuAlaSerAlaArgAlaValProArgAsnValGlnPro 73
1471 CTAATTATCTAGAGAACACAGACATCTGTAATAATATCCATGACATCTTTCATGTT 1530
74 TyrValValTyrGluGluValThrAsnValTrpIleAsnValIleHisPheTyrPro 93
1531 TTTCCCAAGTACAC---GAAGAGGAAATTTGAGTTTATTTTGGCTCTCAATGCAAAACA 1587
94 PheProGlnSerGluGluGluAspGluLeuCysPheLeuArgAlaAsnGluCysIleThr 113
1588 GGTTCCTCGTCATTTATACAAATTTACATCTTATTTAAAGGAAAGCAATATAAAGATCC 1647
114 GlyPheCysHisLeuTyrLysValThrAlaValLeuLysSerGlnGlyTyrAspTrpSer 133
1648 AGTGGTGGCTGCTGCCAAGTATTTCAAGTGTCTTCAAGAGAGAGATAGCAATT 1707
134 GluProPheSerProGly-----GluGlyGluGlnSerLeu 145
1708 ACCAGTGTGATGGAGAGTCTTGGCGCGATGATCTAATCAAGTTGATCAAGTGC 1767
146 ThrAsnAlaIleTrp-----ValAsnGluGlu 154
1768 AGAGGCTGTGATATTTGAAGCACCACCAAGACTCCCTTTTAGAGCATCACCTGTACGTA 1827
155 ThrLysLeuValTyrPheGlnGlyThrLysAspThrProLeuGluHisLeuTyrVal 174
1828 GTCAAGTACGTAATCTCGAGAGGTGACAAAGCTGACTGACCGTGGCTACTCACATCT 1887
175 ValSerTyrGluAlaAlaGlyGluLeuValArgLeuThrThrProGlyPheSerHisSer 194
1888 TGCTGTCAGTCAGTCACTGTGACTCTTTATTAAGTAAGTATAGTAACACAGAAAGATCCA 1947
195 CysSerMetSerGlnAsnPheAspMetPheValSerHisTyrSerSerValSerThrPro 214
1948 CACTGTGTGTCCTTTCAAGCTATCAAGTCTCGAAGTACCAACCACTTGCACAAACAAG 2007
215 ProCysValHisValTyrLysLeuSerGlyProAspAspSerProLeuHisLysGlnPro 234
2008 GAATTTGGGCGACCAATTTGGATTACAGAGTCTCTTCTGACTATATCTCTCCAGAA 2067
235 ArgPheTrpAlaSerMetMetGluAla-----Lys 245
2068 ATTTCTCTTTTGAAGTACTGATTTGATTTGATGATGATGATGATGATGATGATGATGAT 2127
246 IlePheHisPheHisThrArgSerAspValArgLeuTyrGlyMetIleTyrLysProHis 265
2128 GATCTACAGCTGGAAAGAAATATCTACTGTGCTGTTCATATATATGCTGCTCCTCAGGTG 2187
266 AlaLeuGlnProGlyLysLysHisProThrValLeuPheValTyrGlyProGlnVal 285
2188 CAGTTGGTGAATAATTCGGTTTAAAGGAGTCAAGTATTTCCGTTGATATACCTAGCTCT 2247
286 GlnLeuValAsnAsnSerPheLysGlyLysLysTyrLeuArgLeuAsnThrLeuAlaSer 305
2248 CTAGGTTATGCTGTGATGATATACACAGGGGATCTCTCAGCGGGCTTAATTT 2307
306 LeuGlyTyrAlaValValIleAspGlyArgGlySerCysGlnArgGlyLeuArgPhe 325
2308 GAAGGCGCTTTAAATATAAATGGTCAATAGCAATTCAGCATCAGGTGGGAAGACTC 2367
326 GluGlyAlaLeuLysAsnGlnMetGlyGlnValGluLeuGluAspGlnValGluGlyLeu 345
2368 CAATATCTAGCTTCTCGATATGATTTTCAATTCATGATCGTTGGGCATCCACGCTCG 2427
346 GlnPheValAlaIleGlyLysTyrGlyPheLeuAspLeuSerArgValAlaIleHisGlyTrp 365
2428 TCTATGAGGAGTACTCTCCCTGATGCTAATATGAGAGGTGAGATATCTTCAGGTT 2487
366 SerTyrGlyPheLeuSerLeuMetGlyLeuIleHisLysProGlnValPheLysVal 385

2488 GCTATTGCTGGGGCCCCAGTCACCTCTGTGATCTTCTATGATACAGATACACGACCT 2547
386 AlaIleAlaGlyAlaProValThrValTrpMetAlaTyrAspThrGlyTyrThrGluArg 405
2548 TATATGGGTCCACCTGACACAGATGAACAGGGCTATTACTTAGATCTGTGGCCATGC 2607
406 TyrMetAspValProGluAsnAsnGlnHisGlyTyrGluAlaGlySerValAlaLeuHis 425
2608 GCAGAAAAGTTCCCTCTGACCAAAATCGTTTACTGTCTTACATGCTTTCTCGGATGAG 2667
426 ValGluLysLeuProAsnGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGlu 445
2668 AATGTCCATTTTGCACATACACAGTATATCTAGTATTTTGTAGTGGGGCTGGAAAGCCA 2727
446 AsnValHisPhePheHisThrAsnPheLeuValSerGlnLeuIleArgAlaGlyLysPro 465
2728 TATGATTTA-----CAGATCTATCTCTAGGAGAGACAC 2760
466 TyrGlnLeuGlnValAlaLeuProProValSerProGlnIleTyrProAsnGluArgHis 485
2761 ACATAAGAGTTCTGTAATCGGAGAACATTTATGACATCTTTTGGCACTACCTTCAA 2820
486 SerIleArgCysProGluSerGlyGluHisTyrGluValThrLeuLeuHisPheLeuGln 505
2821 GAAAACCTT 2829
506 GluTyrLeu 508
RESULT 8
Q7QBK1_ANOGA
ID Q7QBK1_ANOGA PRELIMINARY; PRT; 886 AA.
AC Q7QBK1;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ENSANGP00000016526 (Fragment).
GN ORFNAMES=ENSANG00000014037;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Culicidae;
OC Anophelinae; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RT "Anopheles gambiae re-annotation."
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAA01008879; EAA08416.2; -; Genomic_DNA.
DR MEROPS; S09.016; -.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001375; Peptidase_S9.
DR InterPro; IPR002469; Peptidase_S9.
DR InterPro; IPR000379; Ser esters.
DR Pfam; PF00930; DppIV_N; 1.
DR Pfam; PF00326; Peptidase_S9; 1.
FT NON TER 886
FT NON TER 886
SQ SEQUENCE 886 AA; 100440 MW; 3C284605CAA57DB4 CRC64;
Alignment Scores: 1.14e-112 Length: 886
Pred. No.: 1

Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
 (7)
 RN NUCLEOTIDE SEQUENCE.
 RP STRAIN=Berkeley.

RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champagne M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
 RA Garza R., Gonzalez M., Guerin H., Kronmiller B., Li P., Liao G.,
 RA Miranda A., Mungall C.J., Nunco J., Pacleb J., Paragas V., Park S.,
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
 RA Celniker S.;
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.

CC -I- INTERACTION:
 CC Q9VZRO:CG12016; NBExp-1; InAct=EBI-105926, EBI-152819;
 DR EMBL; AB003749; AAF56357.2; -; Genomic DNA.
 DR EMBL; BT001499; AAN71254.1; -; mRNA.

DR INACT; Q9VC20; -;
 DR MEROPS; S09.016; -;
 DR Ensembl; CG3744; Drosophila melanogaster.
 DR FlyBase; FBgn0039240; CG3744.

DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0003824; F:catalytic activity; IEA.
 DR GO; GO:0004274; F:dipeptidyl-peptidase IV activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.

DR InterPro; IPR001375; Peptidase_S9.
 DR InterPro; IPR002469; Peptidase_S9B.
 DR InterPro; IPR00930; DPPIV N; I.
 DR Pfam; PF00326; Peptidase_S9; 1.

SQ SEQUENCE 1053 AA; 118053 MW; C94AA663AB464577 CRC64;

Alignment Scores:

Pred. No.: 2e-111 Length: 1053
 Score: 1624.50 Matches: 393
 Percent Similarity: 53.3% Conservative: 159
 Best Local Similarity: 38.0% Wismatches: 330
 Query Match: 29.3% Indels: 153
 DB: 2 Gaps: 30

US-10-825-632-2 (1-3120) x Q9VC20_DROME (1-1053)

QY 14 TCAGAGGCAAGCGCGTGTACTGCGCGCGTCTTCTTATGTCGGCGTTCGCGCGCTG 73
 DB 72 Serf1eThrAenAlaAlaThrThrSerLeuAlaAen-LeuLeuAspGlyPheThrAlaAr 91
 QY 74 GGTGTGTCAGCG 133
 DB 91 gValSerThrAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 111
 QY 134 GCAGCATGAAGCG 187
 DB 111 a-----AlaAlaSerValIleThrAenHisLeuSerSerProThrSe 125
 QY 188 GGGG-----CCGGGGGAAGGAAATGCAACATGGCAGCAATGGAAACAGAACAGC 241
 DB 125 rGlyThrProProHisGlyLeuAepValAepGluGly----- 137
 QY 242 TGGGTGTGAGATATTGAAACTGCGGACTGTGAGAGAAATATTGAATCAGAGATCGCG 301
 DB 138 -----Asp-AspGluCysAspAspCysGluGluAep-----GluAepAep 151
 QY 302 CTAAATGGAGCCTTTTATGTT-----GAGCGGTATCTCGGAGTCAAGCTTA 349
 DB 151 rValAspAenAspGlyHisAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 171
 QY 350 AAAAGTGTGTCGATACAGGAAATATCATGGTACATGATGCTAAGCGCACATG 409
 DB 171 yGlnValValGlnGlnIleArgGlyLysMetCysAenLeuSerSerMetValProThr 191
 QY 410 ATTTATGTTGTGAAGAGGAAATGATCCAGATGAGCTCATTCAGACAGAACTATTACC 469
 DB 191 snValGlnPhe-----ArgHisLeuSerAspGly-----ArgAlaArgCysIlePhe 207
 QY 470 TTGCCATGCTGTGTGAGAACAGAGAAATACACTGTTTATTCTGAAATTCCTCAAACTA 529

DB 207 euGlyThrProGlnSerTrpGlnThrThrLeuLeuPheAlaAepIleAsnLeuThrG 227
 QY 530 TCAATAGAGCAGCAGTCTTAATG----- 552
 DB 227 InSerGluGluGlnGlnLeuLeuValGlnArgLeuGluGlyIleAlaSerAspGluTrpS 247
 QY 552 ----- 552
 DB 247 erProThrMetAsnAlaGlySerProThrSerSerGlyHisGlnProAlaPheLeuPheA 267
 QY 553 -----CTCTCTGGGAAGCCTCTTTGGATCTT-----TTTCAGGCAA 589
 DB 267 snSerLeuProArgProArgLeuProTrpSerProLeuLeuGlnGlnProIleGlnSerS 287
 QY 590 CACTGGACTATGGA-----ATGTATCTTCGAGAAGAAGAACTATTAA 631
 DB 287 erGlyGlySerGlyGlySerGlySerAlaSerProTyrrAlaArgGluTyrrGlnLeuLeuG 307
 QY 632 GAGAAAGAAAACGCATTGGAAACAGTCGGAATTTGCTTACGATTATTCACCAAGGAAGTG 691
 DB 307 InGluArgGlyArgLeuSerThrTrpGlyIleThrSerTyrrGluLeuHisLysProSerG 327
 QY 692 GAACATTTCTGTTTCAAGCCGCTAGTGGAAATTTATCACGTAAAGATGAGGGGCCACAAG 751
 DB 327 lYlYsLeuValPheProCysPheAenAepLeuTyrrGlnCysLeuAep-----ThrG 344
 QY 752 GATTTACGCAACACCTTTAAGGCCCAATCTAGTGGAACACTAGTTGTCCCAAC---ATAC 808
 DB 344 lYtyrAsnSerGlyLeuLeuPheProThrGlnLeuArgThr---CysProGlnTrpThrA 363
 QY 809 GGATGGATCCAAATATTATGCCCGCTGATCCAGACTGGATTGCTTTTATACATAGCAACG 868
 DB 363 laLeuAspProGlnIleCysProGlnAenSerAepMetIleAlaTyrrIleSerAspCysA 383
 QY 869 ATATTGGATATCTAACATCTTACCAGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 928
 DB 383 spLeuPheValThrHisThrLeuSerGlyHisGluLysArgLeuThrTyrrThrSerThrG 403
 QY 929 ACCTGACCAACATGGAAGAGATGCGAGATGCGAGTGGAGTGGTCTTGTCTTCTCAG 988
 DB 403 lYArgHisSerTyrrValAspAlaLeuSerAlaGlyValProSerTyrrValMetGlnG 423
 QY 989 AAGAAATTTGATAGATATTCTGCTATGCTGCTGCTGCCAAAGCTGGAAGCAACTCCCGAG 1048
 DB 423 luGluPheSerArgTyrrGlnGlyPheTrpGlnProHis-----SerA 438
 QY 1049 GTGGTAAATCTTGAAGATCTATATGAAGAAATGATGAATCTGAGTGGAAATTTATTC 1108
 DB 438 snAspGlyIleTyrrArgIleValTyrrGluGluValAepGluSerGluValSerValTy 458
 QY 1109 ATGTTACATCCCTATGTTGAAACAAAGGAGGCGAGATTCATTCCTGTTATCTCTAAACAG 1168
 DB 458 hrPheProSerSerThrAlaMetHisGlyArgValAepGluTyrrArgPheProArgThrG 478
 QY 1169 GTACAGCAATCTTAAGTCACTTTTAAGATGTCAGAAATTAATGATTGATGCTGAAGGAA 1228
 DB 478 lYSerProAenAlaLysSerLysLeuLysLeuValGlnPheValLeuAenGluAlaLeuG 498
 QY 1229 GATCATAGATGTCATAGATAAGCACTTAATCAACCTTTTGAGATTCTA-----TTTG 1282
 DB 498 InValSerGluIleAlaIleLysAepLeu-----ProTyrrSerLeuLeuAlaValPheS 516
 QY 1283 AAGGAGTTGAATATATTCAGAGCTGAGTGGATGCTCTGAGGGGAAAATATGCTTGGTCCA 1342
 DB 516 erTrpLeuGluTyrrIleValArgValGlyTrpThrProAepAlaLysTyrrValTrpValG 536
 QY 1343 TCCTACTAGATCGCTCCAGACTCGCTACAGTAGTGTGATCTACCTACCTGAAATTATTA 1402
 DB 536 InGlyLeuAspArgLysGlnGlnArgLeuAepValIleLeuIleProLeuAenAenPhe- 555
 QY 1403 TCCAGTAGAAGATGATGTTATGGAAGGCGAGACTCATTTGAGTCACTGCTCCCTGAT- 1458


```
Db 556 -----CysGluSerTyrSerSerGlnValSerThrProThrAspSerIleGlyAspHis 574
QY 1459 -----TCTGTGACGCCACTA---ATTATCTATGACAAACAA 1492
Db 574 erTrpArgSerLeuTyrSerArgThrIleThrProLeuGlnValIleTyrThrGluArgS 594
QY 1493 CAGACATCTGGATAAATATCCATGCATCTTCATGTTTCCCAAAAGTCACCAAGAGG 1552
Db 594 erAspSerTrpIleAsnValHisAspMetLeuHis---PheLeuAspLeuThrGluThrS 613
QY 1553 AAATGTAGTTTATTTTTCCTCTGAATGCAAAACAGGTTTCCGTCATTATACAAAATTA 1612
Db 613 erValThrPheLeuTrpAlaSerGlu---GluThrGlyPheArgHisLeuTyrLeuValT 632
QY 1613 CATCTATTTTAAAGGAAGCAAAATATAACGATCCAGTGGTGGCTGCTGCTCA---- 1668
Db 632 hrAlaSerLeuLeuSerGln-----AlaAsnGlyGlnProAspProGlyS 648
QY 1669 -----AGTCATTTCAAGTGTCTCTATCAAGAGG 1696
Db 648 erValGlyAlaGlnProSerPheValAspLeuSerAlaLeuGlnProArgIleLeuAsnL 668
QY 1697 AGATAGCAATTTACAGTGGTGAATGGGAAGTTCTTGGCCGCGCATGCATCTAATATCCAAG 1756
Db 668 ysValAlaLeuThrSerGlyGluTrpGluValLeuAlaArg-----AsnLeuTrpV 685
QY 1757 TTGATGAAGTCAGAAAGCTGTATATTTTGAAGGCCACCAAGACTCCCTTTTAGAGCATC 1816
Db 685 alAspLysAlaAsnLysLeuValTyrPheValGlyLeuArgAspThrProLeuGluLysH 705
QY 1817 ACCTGTACTAGTACAGTAACTCGTAGAGGTCGACAAAGGTCAGTACCGTGGCT 1876
Db 705 isLeuTyrValValSerLeuGluArgProGluHisIleArgLeuLeuThrGluProGlyT 725
QY 1877 ACTCACATTTCTTGC-----TGCATCATGTCAGC 1903
Db 725 yrSerTyrLeuValGluPheAspAspHisPheAsnAspLeuPheProIlePheSerGlnG 745
QY 1904 ACTGTGACTCTTTTAAAGTAAGTATAGTAAC---CAGAAGAATCCACACTGTGTGTCCTC 1960
Db 745 lnCysLysLeuMetLeuLeuValTyrCysAsnIleGlnArgLeuProSerCysLysValM 765
QY 1961 TTTACAGCTATCAAGTCTCGAAGATGCCACCTTGCAAAACAAAGGAATTTTGGGCCA 2020
Db 765 et-----ArgValAsnGlnThrCysSerAsnGlyGlyValAsnGlyI 779
QY 2021 CCATTTTGGATTACAGAGGT-----CCTCTCTCTGACTATATCTC 2059
Db 779 leGlnIleSerLeuValGlyTyrLeuHisGluGlyLysProGluProGlnTyrCys- 798
QY 2060 CTCACAGAATTTTCTCTTTTGAAGTACTACTCGAATTACATTGTATGGATGCTCTACA 2119
Db 799 --ProGlnIlePheSerProGlnLeuProSerGlyAspIleValTyrAlaMetValPheL 818
QY 2120 AGCTTCATGATACAGCTCGGAAGAAATATCCTACTGTCTGTGTTTTCATATATGTTGTC 2179
Db 818 ysProHisAsnPheGluLeuGlyValLysTyrProThrValLeuAsnValTyrGlyGlyP 838
QY 2180 CTCAGGTGCAGTTCGTGAATAATCGGTTTAAAGAGTCAGATTTTCCGCTTGAATACCC 2239
Db 838 roGluValGlnThrValAsnAsnThrPheLysGlyLysHisGlnLeuArgMetHisMetL 858
QY 2240 TAGCTCTCTAGTTTATGTTGTTGATGATAGACAAACAGGGGATCTGTCTACCGAGGCC 2299
Db 858 euAlaGlnGlyTyrCysValIleCysIleAspSerArgGlySerArgHisArgGlyL 878
QY 2300 TTAATTTGAAGGCCCTTTAAATATATAAATGGTCAATAGAAATTCAGCATCAGTGG 2359
Db 878 ysArgPheGluSerHisIleArgGlyArgMetGlyGlnValGluLeuThrAspGlnValA 898
QY 2360 AGGACTCCAATATCTAGCTTCTCGATATGATTTTCACTTACCTAGATCTGTGGGCGCATCC 2419
Db 898 spAlaLeuArgSerLeuSerAspGlnLeuGlyTyrIleAspMetAspArgValAlaIleH 918
```

```
QY 2420 ACGGCTGGTCTCTATGGAGGATACCTCTCCCTGATGGCATTAAATGACAGAGGTACAGATATCT 2479
Db 918 isGlyTyrSpSerTyrGlyGlyTyrLeuSerLeuMetGlyLeuValGlnTyrProLysIleP 938
QY 2480 TCAGGGTGTCTATGTCTGGGGCCCCAGTCACCTCTCTGTGGATCTTCTATGATACAGATACA 2539
Db 938 heLysValAlaIleAlaGlyAlaProValThrAsnTrpGluTyrTyrAspThrGlyTyrT 958
QY 2540 CGGAACGTTATATGGTTCACCTCCCTGACCAAGATGAACAGGGCTATTACTTAGCATCTGGG 2599
Db 958 hrGluArgTyrMetAspMetProGlnAsnAsnGluAlaGlyTyrSerAlaGlySerValL 978
QY 2600 CCATGCAAGCAGAAAGATTTCCCTCTGAAACCAAAATCGTTTACTGCTCTTACATGTTTCC 2659
Db 978 euGluTyrValAsnSerPheProGluGluAspLysArgLeuLeuLeuIleHisGlyLeuI 998
QY 2660 TGGATGAGATGTCCATTTTGCATACCATCAGTATATTTACTGAGTGTCTTTTAGTGGGGCTG 2719
Db 998 leAspGluAsnValHisPheCysHisThrSerArgLeuIleSerAlaLeuAsnLysAlaA 1018
QY 2720 GAAAGCCATATGATTTTACAGATCTATCTCTCAGGAGAGACACAGCATATAGAGTTCCTGAAT 2779
Db 1018 snLysProTyrGluValHisLeuPheProGluGluArgHisSerLeuArgAsnLeuGluS 1038
QY 2780 CGGAGAACATATTGAACACTGCATCTTTTGCACCTACCTTCAA 2820
Db 1038 erAsnLysAsnTyrGluThrLysLeuLeuSerPheLeuGln 1051

RESULT 10
Q9VC19 DROME PRELIMINARY; PRT; 1113 AA.
ID Q9VC19
AC Q9VC19
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE CG3744-PB, isoform B.
GN Name=CG3744; ORFNames=CG3744;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Ballaw R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Buesam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Fabios B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foeller C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hoston D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
```



```
QY 2062 -----CCAGAAATTTCTCTTTTGAAGTACTACTGGATTACATTGATCGG 2109
Db 481 ThrGlnTyrAsnProSerIleHisSerProGlnIleSerSerGlyAspValLeuTyrAla 500
QY 2110 ATGCTCTACAGCCTCATGATCATACAGCCTGGAAAGAAATATCCCTACTGTGCTGTTCATA 2169
Db 501 MetValPheLysProHisAsnPheMetLeuGlyValLysTyrProThrValLeuAsnVal 520
QY 2170 TATGGTGGTCTCTCAGGTGCAGTTGGTGAATAATCGGTTTAAAGGAGTCAAGTATTTCGCG 2229
Db 521 TyrGlyGlyProGlnValGlnThrValSerAsnThrPheLysGlyMetArgGlnLeuArg 540
QY 2230 TTGAATACCTAGCCTCTCTAGTTATGTGGTGTAGTATAGACACACAGGGGATCCTGT 2289
Db 541 MethisMetLeuAlaSerGlnGlyTyrCysValIleCysValAspSerArgGlySerArg 560
QY 2290 CACGAGGGCTTAAATTTGAAGCGCCTTTAAATATATAAATGGTCAAAATAGAAATTCAC 2349
Db 561 HisArgGlyValGluPheGluSerTyrIleArgCysArgMetGlyThrValGluLeuSer 580
QY 2350 GATCAGGTGGAGGACTCCAATATCTAGCTTCTCGATATGATTTCAATTCATCTAGTATCGT 2409
Db 581 AspGlnValGluValLeuArgIleLeuAlaAspGlnLeuGlyTyrIleAspMetAspArg 600
QY 2410 GTGGGCATCCAGCGCTGCTCTCTATGAGGATACCTCTCCCTGATGGCATTAAATGACAGG 2469
Db 601 ValAlaIleHisGlyTyrSerTyrGlyTyrLeuSerLeuMetGlyLeuValGlnTyr 620
QY 2470 TCAGATATCTTCAGGGTGTCTATTGTGGGCCCCAGTCATCTGTGGATCTTCTATGAT 2529
Db 621 ProGluIlePheLysValSerIleAlaGlyAlaProValThrSerTrpGluTyrTyrAsp 640
QY 2530 ACAGATACACGGAACGTTATATGGGTCCACCTGACACAGATGACAGAGGCGCTTACTTA 2589
Db 641 ThrGlyTyrThrGluArgTyrMetAspLeuProAspSerAsnArgSerGlyTyrAlaAla 660
QY 2590 GGATCTGTGGCCATGCACAGCAAGAAAGTTCCCTCTGACCAAAATCGTTTACTGCTCTA 2649
Db 661 GlySerValLeuAsnTyrIleGlnLysPheProAspGluAspAsnArgLeuLeuIle 680
QY 2650 CATGGTCTTCGGATGAGATGTCATTTTGCACATACCATATATTTACTGAGTTTCTT 2709
Db 681 HisGlyLeuIleAspGluAsnValHisPheHisThrSerGlnLeuValSerArgLeu 700
QY 2710 GTGAGGGCTGGAAGCCATATGATTTACAGATCTATCTCAGGAGACACAGCATPAAGA 2769
Db 701 ValArgAlaAsnLysProTyrGlnLeuGlnValTyrProAsnGluArgHisSerLeuArg 720
QY 2770 GTTCTCGAATCGGAGAACATATGAACTGCATCTTTTGCACTACCTTCAAGAAACCTT 2829
Db 721 AsnLeuGluAlaSerLysHisTyrGluThrLysLeuSerPheLeuGlnAsnHisLeu 740

RESULT 12
ID Q7PTT8 ANOGA PRELIMINARY; PRT; 621 AA.
AC Q7PTT8-
DT 01-MAR-2004 (TREMBLrel. 26, Created)
DT 01-MAR-2004 (TREMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE ENSANGP0000015447 (Fragment).
GN ORFNames=ENSANGP0000012958;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Culicidae;
OC Anophelinae; Anopheles.
OX NCBI_TaxID=180454;
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.

[2]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAB01008287; EAA03335.3; -; Genomic_DNA.
DR MEROPS; S09.016; -.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001375; Peptidase_S9.
DR InterPro; IPR002469; Peptidase_S9B.
DR InterPro; IPR000379; Ser_estrs.
DR Pfam; PF00930; DPEIV_N; 1.
DR Pfam; PF00326; Peptidase_S9; 1.
FT NON_TER 1 621
FT TER 621 621
SQ SEQUENCE 621 AA; 70546 MW; F6EA8463A343BBF3 CRC64;

Alignment Scores:
Pred. No.: 5,68e-89 Length: 621
Score: 1321.00 Matches: 279
Percent Similarity: 58.6% Conservative: 99
Best Local Similarity: 43.3% Mismatches: 187
Query Match: 23.8% Indels: 80
Db: 2 Gaps: 13

US-10-825-632-2 (1-3120) x Q7PTT8 ANOGA (1-621)
QY 1054 AAAAAATCTTGAATATCTATATGAAGAAATGATGATCTGAGGTGGAATTAATTAATCATGTT 1113
Db 3 GluValTyrArgIleValTyrGluGluValAspGluSerAspValSerLeuTyrThrPhe 22
QY 1114 ACATCCCTATGTGTGGAACAAGAGGCGCAGATTCATTCCTGTTATCTTAAACAGGTACA 1173
Db 23 ProSerSerGlnSerAlaGlyArgAspTyrGluGluTyrArgPheProArgAlaGlyThr 42
QY 1174 GCAATCTCTAAAGTCACTTTTAAGATGTCAGAAATAATGATGATGCTGGAAGGAGGATC 1233
Db 43 ProAsnAlaLysSerLysLeuLysValGlnPheArgLeuSerGluAsnLeuArgIle 62
QY 1234 ATGATGTCATAGATAGGAACTAATTCACCTTTTGAGATTCATTTTGAAGGAGGTGAA 1293
Db 63 ThrAspValCysIleLysGluLeuGlnCysProLeuThrPheAlaPheProTrpLeuGlu 82
QY 1294 TATATTGCCAGAGCTCGATGGACTCTCTGAGGAAATATGCTTGCTCCATCTACTAGAT 1353
Db 83 TyrIleValArgValGlyTrpThrProAspSerArgTyrValTrpAlaGlnLeuLeuAsp 102
QY 1354 CGCTCCACAGACTCGCCTACAGATAGTGTGTGATCTCACCTGAATATTATTTATCCAGTAGAA 1413
Db 103 ArgProGlnGlnArgLeuGluLeuValLeu-----LeuProVal--- 115
QY 1414 GATGATGTTTGAAGAGGAGGAGACTCATATTGAGTCAGTGCCT- 1455
Db 116 AspAsnPheCysGluIleTyrSerSerSerSerSerSerSerLeuProProArgLysSerSer 135
QY 1456 -----GATTCGTGAGCGCCACTA---ATTATCTATCAAGAA 1498
Db 136 GlyTrpArgSerProLeuAspLysSerThrThrArgProLeuGlnValIleTyrThrGlu 155
QY 1489 ACAACAGACATCTGGATAAATATCCATGACATCTTTCATGTTTTCCTCCCAAGTCACGAA 1548
Db 156 ThrSerSerSerTrpValAsnValHisAspValLeuGln---PheValGluLeuSerGlu 174
QY 1549 GAGGAAATTCAGTTTATTTTTCGCTCTGAATGCAAAACAGAGTTTCCGTCATTATATACAAA 1608
Db 175 GlnGluValThrPheLeuThrAlaSerGlu----GluSerGlyPheArgHisLeuTyrLeu 193
QY 1609 ATTACATCTATTTTAAAGGAAAGCAATATATAACGATCCAGTGGTGGGCTGCCTCTCCA 1668
```

```
Db 194 valThrSerSerLeuSerProAsnGluVal-----SerSerGlyGlyAlaGlyAla--- 210
Qy 1669 AGTCATTTCAAGTGCTCT-----ATCAA 1692
Db 211 ThrAspHisSerLeuProSerMetAlaCysIleGlySerThrLeuValAlaAargIleVal 230
Qy 1693 GAGGAGATAGCATACCAAGTGTGAATGGAGTCTTGGCGGCATGGATCTAATATC 1752
Db 231 GlnLysValThrLeuThrGlyGlyAspTrpGluValLeuGlyArg-----AsnVal 247
Qy 1753 CAAGTGTGAAGTCAGAGGCTGGTATTTTGAAGGCCAAAGACTCCCTTTAGAG 1812
Db 248 TrpTrpAspArgValArgGlnLeuValTyPheMetGlyLeuAargIleThrProLeuGlu 267
Qy 1813 CATCACCTGTAGTGTAGTACGTAAATCTCGGAGAGGTGCACAAAGCTGACTGACCGT 1872
Db 268 LysHisLeuTyValValSerLeuAlaGlnProAsnGlnLeuAargLeuLeuThrMetPro 287
Qy 1873 GGCTACTCACATCTTCTGCTGCATCAGTACGACTGTGACTTCTTTAAGTAAGTATAGT 1932
Db 288 GlyTySerPheThrValGluPheAsnAspAspCysThrLeuPheLeuGlnThrTyCys 307
Qy 1933 AACCCAGGAAGATCCACACTGTGTGCTCCTTTTACAGCTA-----TCAAGTCCCTGAA 1983
Db 308 AsnIleSerThrLeuProSerTrpGluLeuValArgIleAlaHisAspSerAsnThrAla 327
Qy 1984 GATGACCCAACTTGCANAAACAAGGAATTTGGCCACCATTTCGATTTCAGCAGGTCT 2043
Db 328 AsnGlyAsnGlyCys-----SerHisGlyPro 336
Qy 2044 CTTCTGACTACTCTCT----- 2061
Db 337 ThrProProThrProIleAspAlaLeuAargLeuCysSerValGlyTyLeuThrGlu 356
Qy 2062 -----CAGAAATTTCTCTTTTGAAGTACTACT 2091
Db 357 GlyGlyProSerGluAsnThrGlnTyAsnProSerIleHisSerProGlnIleSerSer 376
Qy 2092 GGATTTACATGTATGGAGTGTCTACAGCTCATGATCTACAGCTCGGAGGAAGAAATAT 2151
Db 377 GlyAspValLeuTyAlaMetValPheLysProHisAsnPheMetLeuGlyValLysTy 396
Qy 2152 CCTACTGTGCTGTTCATATATGTGTGTCTCTCAGTGTGCTGTGTAATTCGGTTTAA 2211
Db 397 ProThrValLeuAsnValTyGlyGlyProGluValGlnThrValSerAsnThrPheLys 416
Qy 2212 GGAGTCAAGTATTTCCGCTTGAATACCTAGCTCTCTAGGTATGTGTGTGTGTAGTGATA 2271
Db 417 GlyMetArgGlnLeuArgMetHisMetLeuAlaSerGlnGlyTyCysValIleCysVal 436
Qy 2272 GACAAAGGGGATCCTGTCCAGGAGGCTTAAATTTGAAGGCGCTTTAAATATATAAATG 2331
Db 437 AspSerArgGlySerArgHisArgGlyValGluPheGluSerTyIleArgArgMet 456
Qy 2332 GGTCAATAGAAATTTGACATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATGAT 2391
Db 457 GlyThrValGluLeuSerAspGlnValGluValLeuAargIleLeuAlaAspGlnLeuGly 476
Qy 2392 TTCATTACATAGTATCGTGGCCATCCAGCTGCTCTATGAGGAGTACTCTCCCTG 2451
Db 477 TyIleAspMetAspArgValAlaIleHisGlyTrpSerTyGlyTyLeuSerLeu 496
Qy 2452 ATGCATTAATGACAGGTGATGATCTTCAGGTTGCTATGCTGGGCGCCGCTACT 2511
Db 497 MetGlyLeuValGlnTyProGluIlePheLysValSerIleAlaGlyAlaProValThr 516
Qy 2512 CTGTGGATCTCTATGATACGGATACACGGACGCTTATATGGGTCACTCCCTGACAGAT 2571
Db 517 SerTrpGluTyTyAspThrGlyTyThrGluArgTyPheMetAspLeuProAspSerAsn 536
Qy 2572 GAACAGGGCTATTACTTAGATCTGTGGCCATCCAGACAGAAAGTTTCCCTCTCAACCA 2631
Db 537 ArgSerGlyTyTyAlaAlaGlySerValLeuAsnTyIleGlnLysPheProAspGluAsp 556
Qy 2632 ATCTCTTACTGCTCTTACATGTTTCTCGATGAGAAATCTCCATTTTGACATACCAGT 2691
Db 557 AsnArgLeuLeuIleIleHisGlyLeuIleAspGluAsnValHisPheHisHisThrSer 576
Qy 2692 ATATTACTAGCTTTTATTAGTGGGCTGGAAAGCCATATGATTACAGATCTATCTCTCAG 2751
Db 577 GlnLeuValSerArgLeuValArgAlaAsnLysProTyGlnLeuGlnValTyProAsn 596
Qy 2752 GAGACACACAGATTAAGTCTCTGAATCGGAGAACATTAATGATCACTCTTTTGCAC 2811
Db 597 GluArgHisSerLeuArgAsnLeuGluAlaSerLysHisTyGlnThrLysLeuLeuSer 616
Qy 2812 TACCTTCAAGAAAC 2826
Db 617 PheLeuGlnAsnHis 621
RESULT 13
Q5TXJ2 ANOGA
ID Q5TXJ2 ANOGA PRELIMINARY; PRT; 557 AA.
AC Q5TXJ2;
DT 01-FEB-2005 (TREMBLrel. 29, Created)
DT 01-FEB-2005 (TREMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TREMBLrel. 29, Last annotation update)
DE ENSANGP0000029249 (Fragment).
GN ORFNames=ENSANG0000012958;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Culicidae;
OC Anophelinae; Anopheles.
OX NCBI_TaxID=180454;
RN (1)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RT "Anopheles gambiae re-annotation."
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN (2)
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RT Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
RL -1- CAUTION: the sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAA01008287; EAL42106.1; -; Genomic_DNA.
DR GO; GO:0003824; P: catalytic activity; IEA.
DR GO; GO:0008236; P: serine-type peptidase activity; IEA.
DR GO; GO:0006508; P: proteolysis and peptidolysis; IEA.
DR InterPro; IPR001375; Peptidase_S9.
DR InterPro; IPR002469; Peptidase_S9B.
DR InterPro; IPR000379; Ser esters.
DR Pfam; PF00930; DPPIV_N; I.
DR Pfam; PF00326; Peptidase_S9; 1.
FT NON_TER 1 557
SQ SEQUENCE 557 AA; 62685 MW; 08AAA0944ACD0808 CRC64;
Alignment Scores:
Pred. No.: 9,64e-76 Length: 557
Score: 1142.50 Matches: 253
Percent Similarity: 57.7% Conservative: 84
Best Local Similarity: 43.3% Mismatches: 198
Query Match: 20.6% Indels: 49
DB: 2 Gaps: 12
US-10-825-632-2 (1-3120) x Q5TXJ2_ANOGA (1-557)
Qy 1063 AGAATCTATATGAGAAATGATGAATCTGAGTGGAAATATTATTCATGTTACATCCCT 1122
Db 1 ArgIleValTyGluGluValAspGluSerAspValSerLeuTyThrPheProSerSer 20
```

```
QY 1123 ATGTTGGAACAAGAGGCGCAGATTCAATCCGTTATCTCTAAACAGGTACAGCAAAATCTCT 1182
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
QY 21  GlnSerAlaGlyArgAspTyrGluGluTyrArgPheProArgAlaGlyThrProAsnAla 40
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
QY 1183 AAAGTCACATTTAAGATGTCAGAAATAATGATGCTGCTGAAGAGGATCATAGATCTC 1242
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
QY 41  LysSerLysLeuLysLeuValGlnPheArgLeuSerGluAsnLeuArgIleThrAspVal 60
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
QY 1243 ATAGATAAGGAACAAATCAACCTTTTGAGATTCATTGTAAGAGGTGAATATATTGCC 1302
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
QY 61  CysIleLysGluLeuGlnCysProLeuThrPheAlaPheProTrpLeuGluTyrIleVal 80
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
QY 1303 AGAGCTGGATGACCTCCTCAGGGAATAATGCTTGGTCCATCTCTACTAGATCCTCCAG 1362
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
QY 81  ArgValGlyTrpThrProAspSerArgTyrValTrpAlaGlnLeuLeuAspArgProGln 100
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
QY 1363 ACTCGCTCACGATAGTGTGATCTCACTCACTGAAATATTATTATCCAGTAGAAGATGATGT 1422
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
QY 101  GlnArgLeuGluLeuValLeuLeuProValAspAsnPheCysSerThrAla-----Ala 118
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
QY 1423 ATGGAAGGACAGACATCAATGAGTCAGTGCCTGATCTGTGACGCCCACTAATATCTAT 1482
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
QY 119  ArgArgArgProThrAlaGlyArgProProGlyAlaMetGlyProGlyIleTyr 138
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
QY 1483 GAAGAAACAACAGACATCTGGATAAATATCCATGACATCTTTCATGTTTTCCTCCCAAGT 1542
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
QY 139  ThrGluThrSerSerTrpValAsnValHisAspValLeuGln---PheValGluLeu 157
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
QY 1543 CACGAAGAGAAATGAGTTATTTTTCCTCTGANTGCCAAACAGAGTTTCCGTCAATTTA 1602
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
QY 158  SerGluGlnGluVal-----ThrThrAsnSerHisPro 168
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
QY 1603 TACAAATTTACATCTATTTTAAAGCAAGCAAAATAT-----AAACGATCCAGTGTGGG 1656
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
QY 169  ThrAlaHisProLysGlnGlnArgGluGlyAlaHisGlyThrGluValSerSerGlyGly 188
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
QY 1657 CTGCTGCTCCCAAGTGATTTCAAGTGCTCT----- 1686
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
QY 189  AlaGlyAla---ThrAspHisSerLeuProSerMetAlaCysIleGlySerThrLeuVal 207
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
QY 1687 -----ATCAAGAGGAGATAGCAATTACCAGTGGTGAATGGGAAGTTCTTGCCCGCAT 1740
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
QY 208  AlaArgIleValGlnLysValThrLeuThrGlyGlyAspTrpGluValLeuGlyArg--- 226
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
QY 1741 GGATCTAATATCCAAAGTTGATGAGTCAGAGGCTGGTATATTTTGAAGCAGCAACAAGAC 1800
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
QY 1801 TCCCTTTTAGAGCATCACCCTGACGTAGTACGTAGTAAATCTCGAGAGGTGACACAGG 1860
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
QY 245  ThrProLeuGluLysHisLeuTyrValValSerLeuAlaGlnProAsnGlnLeuArgLeu 264
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
QY 1861 CTGACTGACGGTGGTACTACATCTTGTGTCATCATGCTGAGCTGAGCTGCTCTTT--- 1917
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
QY 265  LeuThrMetProGlyTyrSerPheThrVal-----GluPheAsnAspLeuValArg 281
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
QY 1918 ATAAGTAAGTATAGTAACAGAGAATCCACACTGTGTGTCCTTTTACAAGCTCAAGT 1977
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
QY 282  IleAlaHisAspSerAsnThrAlaAsnGlyAsnGlyCysSer-----HisGly 297
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
QY 1978 CTGGAAGATGACCCCACTTGCACAAACAAAGGAATTTTGGGCCACCACTTTTGGATTCAGCA 2037
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
QY 298  ProThrProProThr-----ProIleAspAlaLeuArgLeuCysSerVal 313
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
QY 2038 GGTCTCTTCTCCTGATATATCTCT-----CCAGAAATTTCTCT 2076
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
QY 314  GlyTyrLeuThrGluGlyGlyProSerGluAsnThrGlnTyrAsnProSerIleHisSer 333
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
QY 2077 TTTGAAAGTACTACTGGATTTACATTTGATGGATGCTCTACAAGCCTCATGATCTACAG 2136
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
QY 334  ProGlnIleSerSerGlyAspValLeuTyrAlaMetValPheLysProHisAsnPheMet 353
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
QY 2137 CTTGGAAGAAATAATCTACTGTGCTGTTTCATATATGTTGCTCTCAGGTGCAGTGTG 2196
```

```
Db 354  LeuGlyValLysTyrProThrValLeuAsnValTyrGlyGlyProGluValGlnThrVal 373
QY 2197 AATAATCGGTTTAAAGAGCTCAAGTATTTCGCTTGAATACCTAGCTCTCTAGGTTAT 2256
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
QY 374  SerAsnThrPheLysGlyMetArgGlnLeuArgMetHisMetLeuAlaSerGlnGlyTyr 393
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
QY 2257 GTGTTGTGTAGTATAGACAAACAGGGGATCTGTCTACCGAGGGCTTAAATTTTGAAGGCGCC 2316
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
QY 394  CysValIleCysValAspSerArgGlySerArgHisArgGlyValGluGluPheGluSerTyr 413
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
QY 2317 TTTAAATATATAAATGGGTCAATAGAAATTTGACGATAGGTGGAGGACTCCAATATCTA 2376
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
QY 414  IleArgArgArgMetGlyThrValGluLeuSerAspGlnValGluValLeuArgIleLeu 433
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
QY 2377 GCTTCTCGATATGATTTTCACTGACTAGATCTGTGGGCATCCAGGCTGTCTCTATGGA 2436
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
QY 434  AlaAspGlnLeuGlyTyrIleAspMetAspArgValAlaIleHisGlyTrpSerTyrGly 453
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
QY 2437 GGATACCTCTCCTCGATGCGCATTAATGTCAGAGGTTCAGATATCTTCAGGGTGTCTATTGCT 2496
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
QY 454  GlyTyrLeuSerLeuMetGlyLeuValGlnTyrProGluIlePheLysValSerIleAla 473
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
QY 2497 GGGGCCCCAGCTCACTCTGTGGATCTTCTATGATACAGATACACGGAACGTTATATGGT 2556
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
QY 474  GlyAlaProValThrSerTrpGluTyrAspThrGlyTyrThrGluArgTyrMetAsp 493
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
QY 2557 CACCTGACACAGAAATCAACAGGCTATTACTTAGGATCTGTGGCATCGAAGCAAGAAAG 2616
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
QY 494  LeuProAspSerAsnArgSerGlyTyrAlaAlaGlySerValLeuAsnTyrIleGlnLys 513
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
QY 2617 TTCCCTCTGAACAAATCGTTTACTGCTCTTACATGTTCTTCCTGGATGAGAATGTCCAT 2676
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
QY 514  PheProAspGluAspAsnArgLeuLeuIleHisGlyLeuIleAspGluAsnValHis 533
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
QY 2677 TTGCGCATACACAGTATATATCTGAGTTTCTAGTGGTGGTGGAGGCTGGAAGCCATATGTTA 2736
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
QY 534  PheHisHisThrSerGlnLeuValSerArgLeuValArgAlaAsnLysProTyrGlnLeu 553
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
QY 2737 CAGATCTATCT 2748
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||
QY 554  GlnValTyrPro 557
Db      |||      |||      |||      |||      |||      |||      |||      |||      |||

RESULT 14
Q54U01.DICDI PRELIMINARY; PRT; 803 AA.
ID Q54U01.DICDI
AC Q54U01;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DE Hypothetical protein.
GN ORFNames=DD80205566;
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX NCBI_taxID=44689;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AX4;
RA Eichinger L., Pachbat J.A., Gloeckner G., Rajandream M.-A.,
RA Sugeng B., Berriman M., Song J., Olsen R., Szafranski K., Xu Q.,
RA Tunggal R., Kummerfeld S., Madera M., Konfortov B.A., Rivero F.,
RA Bankier A.T., Lehmann R., Hamlin N., Davies R., Gaudet P., Fey P.,
RA Pilcher K., Chen G., Saunders D., Sodergren E., Davis P.,
RA Kerhornou A., Nle X., Hall N., Anjard C., Hemphill L., Bason N.,
RA Farbrother P., Desany B., Just E., Morio T., Rost R., Churcher C.,
RA Cooper J., Haydock S., van Driessche N., Cronin A., Goodhead I.,
RA Muzny D., Mourier T., Pain A., Lu M., Harper D., Linday R.,
RA Hauser H., James K., Quiles M., Mohan M.B., Saito T., Buchrieser C.,
RA Wardrop A., Felder M., Thangavelu M., Johnson D., Knights A.,
RA Lourishghara H., Hernandez J., Rabinowitsch E., Steffen D., Sanders M.,
RA Urushihara H., Kohara Y., Sharp S., Simmonds M., Spiegler S., Tivey A.,
RA Sugano S., White B., Walker D., Woodward J., Winckler T., Tanaka Y.,
```


RA Shauleky G., Schleicher M., Weinstock G., Rosenthal A., Cox E.C.,
RA Chisholm R.L., Gibbs R., Loomis W.P., Platzner M., Kay R.R.,
RA Williams J., Dear P.H., Noegel A.A., Barrell B., Kuspa A.,
RT "the genome of the social amoeba Dictyostellium discoideum";
RL Nature 0:0-0(2005).
CC -I- CAUTION: The sequence shown here is derived from an
EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
preliminary data.
CC EMBL; AA01000074; EAL66689.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 803 AA; 93187 MW; 33846B370C039FAE CRC64;

Alignment Scores:
Pred. No.: 2,31e-74 Length: 803
Score: 1124.50 Matches: 291
Percent Similarity: 48.7% Conservative: 146
Best Local Similarity: 32.4% Mismatches: 310
Query Match: 20.3% Indels: 151
Gaps: 31

US-10-825-632-2 (1-3120) x Q54U01_DICDI (1-803)

QY 283 ATTGAATCAGAGATCGGCTAAATGGAGCCCTTTATGTTGAGCGGTATTTCTGGAGT 342
Db 8 IlegluGlyArgsp-----GluleuLeuaspysglu 18

QY 343 CAGCTTAAAGCTGCTGGCGATACAGAAATATCATCGGTACATGCTGCTAAGCA 402
Db 19 GluleuLeuaspysgluGlyGluSerGly-----SerSerAla 32

QY 403 CCACATGATTTTC-----ATGTTGTGAAGGATGATCCA 438
Db 33 ProTyrSerPheGluPheAsnAspSerLysAsnHisLeuTyrPheLeuSerAsnVallys 52

QY 439 GATGACCTTATTCAGACAGAAATCTATTACCTTCCCATGCTGTTGGTGAAGCAAGAAAT 498
Db 53 AspGlnLysAsnHisLeuAsnHisLeuTyrPheMetAsnSerLysGluLeu 72

QY 499 ACACGTGTTTATTCGAAATCCCAAACTATCAATAGACGACGACGCTTATGCTCTCT 558
Db 73 LysProLeuPheAsnTyrValAspLysValGlu----- 84

QY 559 TGGAGCCCTTTTGGATCTTTTCAGGCAACACTGGACTATGGAATGATTTCTCGAGAA 618
Db 85 -----LeuSerIleGlu 88

QY 619 GAAGAACTATTAGAGAAAGAAACGCAATTGGACAGCTCGGAATTCCTTACGATTAT 678
Db 89 AspGlnLeuGlnArgGluArgMetArgThrAlaAlaAsnGlyIleThrGlnPheThrPhe 108

QY 679 CACCAAGGAGTGGAAACATTTCTGTTTCAAGCCGTAGTGGAAATTTATCACGTAAGAAT 738
Db 109 AspGlnLysHisGlnPheIleIleAlaProIleAsnAsnLysIleAsnLysIleAspIle 128

QY 739 GGAGGCCCAAGGATTTACGCAACACCTTTTAAGGCCCAATCTAGTGGAAACTAGTTGT 798
Db 129 -----LysGluSerIleThrLysProIleLysGluLeuValGlyGluThrTyrAsn 146

QY 799 CCCACATACGGATCGATCCAAATATTATGCCCGCTGATCCAGCTGGATGCTGTTTATA 858
Db 147 HisGlnIleSerAlaAspGlyLysIle-----ValSerPheLeu 159

QY 859 CATAGCAACGATATTGGATATCTAACATCGTAACACAGAGAAGAAAGGACACTCAT 918
Db 160 LysAspLysAspIleThrPheThrAspIleSerThrAsnAlaMetTyrArgIleThrPhe 179

QY 919 GTGCACAAATGAGTACCAACATGGAAGAAGATGCCAGCTGAGTGGAGTCGTACTTT 978
Db 180 SerAsnAspGlu-----LysHisLysPheArgTyrAlaGlyAspIleGlyPhe 196

QY 979 GTTCTCCAGAGAAATTTGATAGATATTCGCTATTGGTGTGCTCCAAAGCTGAAACA 1038
Db 197 IleTyrAlaGluGluPheSerArgTyrThrGlyTyrTrpSerProIleValGlyThr 216

QY 1039 ACTCCAGCTGGTGGTAAA---ATTCTTGAATTTCTATATGTAAGAAAATGATGAATCTGAG 1095
Db 217 CysValLysThrGlyLysProMetTyrThrIleCysTyrLeuGluGluAspGluThrAsn 236

QY 1096 GTGGAAATATTATCATGTTACATCCCTATGTTGGAAACAAGG---AGGCGACATTCATTC 1152
Db 237 ValMetAspTyrHisIle-----ProThrSerAspLeuArgGlyLysThrThrGlnTyr 254

QY 1153 GCTTATCTAAACACAGGTACAGCAAT----- 1179
Db 255 LysTyrProLeuAlaGlyGlyLysAsnSerIleCysLysValCysLeuValSerPheVal 274

QY 1180 -----CCTAAAGTCACTTTTAAGATGTCAGAAATAATGATTTGATGCTGAAGGAAGATC 1233
Db 275 LeuProThrArgThrThrPheGlnAspSerLys----- 285

QY 1234 ATAGATGTCATAGATAAGAACTAAATTCACCTTTTGGATTTCTATTTGAAGAGCTGAA 1293
Db 286 IleGluIleValLysSerGluLeuPheAsp---LeuLysThrGlnPheProTrpAlaGlu 304

QY 1294 TATATTGCCAGCTGGATCGACTCTCGAGGAAATATGCTTGGTCCATCTCTACTAGT 1353
Db 305 TyrIleThrArgAlaGlyTrpThrProAsnGlyHisSerIleTyrLeuGlnLeuLeuAsp 324

QY 1354 CGCTCCACAGACTCGCCTACAGATAGTGTTCATCTCACCTGAAATATTATTCACAGTAGAA 1413
Db 325 ArgLysGlnGlnHisLeuAlaLeuValMetValProLeuHisValPhe-----AlaGlu 342

QY 1414 GATGATGTTATGGAAGGACAGACTCATTTGAGTGCCTGATTCGTGTGAGCGCACATA 1473
Db 343 AspTyrSerSerSerSerSerSerSerSerValLysSerIleProLys-----LeuProVal 360

QY 1474 ATTATCTATGAAGAAACAACAGACATCTGGATAAATAATCCATGCATCTTTTCATGTTT 1533
Db 361 LeuIle---GluGluThrThrSerValTrpIleAsnIleGluPheSerPheGln---Phe 378

QY 1534 CCCCAAGTCAAGAGAGAAATTCAGTTTATTTTGCCTCTGAATGCACAAACAGAGTTTC 1593
Db 379 LeuLysSerIleGluAsnGlnLeu-----IleTrpSerAsnGlu---GlnSerGlyTyr 395

QY 1594 CQTATATTATACAAATATACATCTATTTTAAAGAAAGCAATATAAATGATCCAGTGGT 1653
Db 396 ArgHisLeuTyr-----LeuIleLysTrpAspLysAsnPheThrAsnIleGln 411

QY 1654 GGGCTGCTGCTCCAAAGTGATTTCAAGTGCTCTATCAAGAGGAGTAGCAATACCACT 1713
Db 412 SerThrProIleThrLeuSer---ThrCysAsnAspAsnAspAsn-----AspAsn 427

QY 1714 GGTGAATGGAAAGTTCTTTGGCCGCGATGGATCTAATATCCAGTTGATGAAGTCAGAGG 1773
Db 428 AsnAsnTrpMetVal-----SerSerAspAspIleHisIleAspGluLysArgLys 444

QY 1774 CTGCTATATTTTGAAGGACCAAGACTCCCTTTTAGAGCATCACCTGTACGTAGTCAGT 1833
Db 445 LeuValTyrPheThrGlyThrLysAspThrCysLeuGluGlnHisLeuTyrValThrArg 464

QY 1834 TAGCTAAATCCT---GGAGAGGTGACAAAGGCTGACTGACCGCTGCTACTCACAT---TCT 1887
Db 465 PheAspLysProAsnSerGluIleLysArgLeuSerHisAlaAsnPheSerHisArgSer 484

QY 1888 TGTGCTCAGTCAGCAGCTGCTGCTCTTTTATAGTAGTATAGTAAACCAAGAGAAAT--- 1944
Db 485 IleSerIleSerSerAsnPheLysPheIleThrThrTyrSerAsnIleSerThrIle 504

QY 1944 ----- 1944

QY 505 SerLysThrGluValPheAspLeuIleTyrAsnAspAspAsnAsnAspAsnAspIle 524

QY 1945 ---CCACACTGTGTGCTCTTTTACAGCTATCAAGTCTGAGATGATGCCCACTTGCAAA 2001

QY 525 TyrProIleValLysSerSerPhePheIleAsnAspAspAspAspAspAspAspAspAsp 544

Db				:::						353
	Trp	Trp	Cys	Pro	Thr	Ala	--	Ser	Trp	Glu
	Glu	Glu	Glu	Glu	Glu	Glu	Glu	Glu	Glu	Glu
Qy	1075	GA	GA	AA	AA	AT	GA	AG	GT	GA
		AT	GA	AT	CT	GA	AG	GT	GA	AA
		AT	GA	AT	CT	GA	AG	GT	GA	AA
Db	354	Glu	Glu	Val	Ala	Pro	Ser	Pro	Ala	Leu
		Glu	Glu	Val	Ala	Pro	Ser	Pro	Ala	Leu
		Glu	Glu	Val	Ala	Pro	Ser	Pro	Ala	Leu
Qy	1135	AG	AG	GG	CG	AT	TC	CT	TA	AA
		AG	AG	GG	CG	AT	TC	CT	TA	AA
		AG	AG	GG	CG	AT	TC	CT	TA	AA
Db	374	Arg	Gly	Thr	Ala	Pro	Ser	Pro	Ala	Leu
		Arg	Gly	Thr	Ala	Pro	Ser	Pro	Ala	Leu
		Arg	Gly	Thr	Ala	Pro	Ser	Pro	Ala	Leu
Qy	1195	AA	GA	TC	CA	GA	AA	TA	AT	GA
		AA	GA	TC	CA	GA	AA	TA	AT	GA
		AA	GA	TC	CA	GA	AA	TA	AT	GA
Db	394	Lys	Leu	Ala	Glu	phe	Gln	Thr	Ala	Glu
		Lys	Leu	Ala	Glu	phe	Gln	Thr	Ala	Glu
		Lys	Leu	Ala	Glu	phe	Gln	Thr	Ala	Glu
Qy	1255	CT	AA	TT	CA	CC	TT	TT	GA	CT
		CT	AA	TT	CA	CC	TT	TT	GA	CT
		CT	AA	TT	CA	CC	TT	TT	GA	CT
Db	414	Leu	Val	Gln	Pro	Phe	Ser	Leu	Phe	Pro
		Leu	Val	Gln	Pro	Phe	Ser	Leu	Phe	Pro
		Leu	Val	Gln	Pro	Phe	Ser	Leu	Phe	Pro

Search completed: April 14, 2006, 11:31:04
 Job time : 779 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model
Run on: April 14, 2006, 11:31:23 ; Search time 18.6 Seconds
(without alignments)
2773.636 Million cell updates/sec

Title: US-10-825-632-2
Perfect score: 5552
Sequence: 1 aagtgctaaagctccgagg.....agaattactaaaaaaaaa 3120

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 572060 seqs, 82675679 residues
Total number of hits satisfying chosen parameters: 1144120

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Maximum Match 100%
Listing first 45 summaries

Post-processing: Minimum Match 0%

Command line parameters:
-MODEL=frame_n2p.model -DEV=xlp
-O=/abs/ABSSWEB_spool/US10825632/runat_14042006_092353_3529/app_query.fasta_1
-DB=Issued Patents AA -QFMT=fastan -SUFFIX=n2p.ra1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=PCT -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abs02p -USER=US10825632 @CIGN 1.1.71 @runat_14042006_092353_3529 -NCPU=6
-ICPU=3 -NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5 COMB.pcp:*
2: /cgn2_6/ptodata/1/iaa/6 COMB.pcp:*
3: /cgn2_6/ptodata/1/iaa/H COMB.pcp:*
4: /cgn2_6/ptodata/1/iaa/PCUTS COMB.pcp:*
5: /cgn2_6/ptodata/1/iaa/RE COMB.pcp:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pcp:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4700	84.7	882	2	US-09-976-674-1
2	4700	84.7	882	2	US-10-070-464-1
3	3607.5	65.0	690	2	US-09-976-674-7
4	3513.5	63.3	661	2	US-09-976-674-11
5	3504	63.1	658	2	US-09-976-674-19
6	3236	58.3	613	2	US-09-976-674-21
7	2871	51.7	892	2	US-09-976-674-23
8	2871	51.7	892	2	US-09-976-674-27
9	2870	51.7	863	2	US-09-976-674-3
10	2821.5	50.8	879	2	US-09-976-674-33
11	2821.5	50.8	879	2	US-09-976-674-35
12	2422	43.6	465	2	US-10-070-464-5

13	2407	43.4	832	2	US-09-976-674-29	Sequence 29, Appl
14	2407	43.4	832	2	US-09-976-674-31	Sequence 31, Appl
15	2357.5	42.5	819	2	US-09-976-674-37	Sequence 37, Appl
16	2357.5	42.5	819	2	US-09-976-674-39	Sequence 39, Appl
17	1836.5	33.1	360	2	US-10-070-464-7	Sequence 7, Appl
18	1808	32.6	358	2	US-09-976-674-13	Sequence 13, Appl
19	1645.5	29.6	310	2	US-09-794-236-4	Sequence 4, Appl
20	1645.5	29.6	310	2	US-10-070-464-3	Sequence 3, Appl
21	1392	25.1	518	2	US-09-976-674-25	Sequence 25, Appl
22	1278	23.0	241	2	US-09-976-674-9	Sequence 9, Appl
23	1007.5	18.1	194	2	US-09-976-674-17	Sequence 17, Appl
24	724.5	13.0	981	2	US-09-902-540-16812	Sequence 16812, A
25	616	11.1	710	2	US-09-518-550-28	Sequence 28, Appl
26	582.5	10.5	771	2	US-09-462-284-2	Sequence 2, Appl
27	582.5	10.5	771	2	US-09-079-592-2	Sequence 29, Appl
28	572.5	10.3	732	2	US-09-518-550-29	Sequence 30, Appl
29	572	10.3	732	2	US-09-518-550-30	Sequence 42, Appl
30	569.5	10.3	676	2	US-09-518-550-42	Sequence 6, Appl
31	529	9.5	766	2	US-10-002-593-6	Sequence 6146, Ap
32	529	9.5	766	2	US-09-949-016-6146	Sequence 6, Appl
33	529	9.5	766	2	US-10-423-714-6	Sequence 3, Appl
34	525	9.5	755	4	PCT-US93-07923-3	Sequence 3, Appl
35	525	9.5	759	4	PCT-US93-07923-2	Sequence 3, Appl
36	525	9.5	766	1	US-08-230-491A-3	Sequence 3, Appl
37	525	9.5	766	1	US-08-619-280A-3	Sequence 3, Appl
38	525	9.5	766	1	US-08-940-391-3	Sequence 3, Appl
39	525	9.5	766	2	US-09-794-236-1	Sequence 1, Appl
40	525	9.5	766	2	US-09-265-606-3	Sequence 1, Appl
41	525	9.5	775	2	US-09-949-016-10450	Sequence 10450, A
42	519	9.3	766	2	US-09-518-550-27	Sequence 27, Appl
43	488.5	8.8	737	2	US-09-902-540-11421	Sequence 11421, A
44	484.5	8.7	818	2	US-09-462-845-3	Sequence 3, Appl
45	484.5	8.7	818	2	US-10-402-312-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-09-976-674-1
; Sequence 1, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-1

Alignment Scores:
Pred. No.: 0
Score: 4700.00
Percent Similarity: 100.0%
Best Local Similarity: 100.0%
Query Match: 84.7%
DB: 2
Length: 882
Matches: 882
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0
US-10-825-632-2 (1-3120) x US-09-976-674-1 (1-882)

QY 214 ATGCGACGACGAATGGAACAGACAGCTGGGTGTGAGATATTTGAACCTCGGACGTGT 273
|||||
Db 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
|||||

QY	274	GAGGAGAAATATTGAATCAGAGTACGGCTAAATTTGGAGCCCTTTTATATGTTGACGGTAT	333
Db	21	GlulAuenileGluSerGlnAepAsgProLysLeuGluProPheTyValGluArgTyr	40
QY	334	TCCTGGAGTCAGCTTAAAGAGCTGCTTGGCGATACAGAAAATATCATGGCTACATGATG	393
Db	41	SerTrpSerGlnLeuLysLysLeuLeuAlaAepThrArgLysTyHisGlyTyMetMet	60
QY	394	GCTAAGGCCACCATCATTTTCATGTTTGTGAAGAGAAATGATCCAGATGGACCTCAVTTCA	453
Db	61	AlaLysAlaProHisAepPheMetPheValLysArgAenAepProAepGlyProHisSer	80
QY	454	GACGAATCTATTACCTCTGCCATGCTCGTGAGACACAGAAAATACACTGTTTATTCT	513
Db	81	AspArgIleTyrIleuAlaMetSerGlyGluAenArgGluAenThrLeuPheTySer	100
QY	514	GAATTCCCAAAATCATTAATAGACAGACAGCTTTAATGCTCTCTTGGAGCCCTTTTG	573
Db	101	GluIleProLysThrIleAenArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu	120
QY	574	GATCTTTTTCAGGCAACACTGGACTATGAAATGTCTTCTTACGATTATCAACAAGGAGTGA	633
Db	121	AspLeuPheGlnAlaThrLeuAepTyrGlyMetTyrSerArgGluGluGluLeuLeuArg	140
QY	634	GAAGAAGAACGCAATCGAACAGTCGGAATGTCTTCTTACGATTATCAACAAGGAGTGA	693
Db	141	GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly	160
QY	694	ACATTTCTGTTTCAAGCCGGTAGTGGAAATTTATCAAGTAAAGATGGAGGGCCCAAGGA	753
Db	161	ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAepGlyGlyProGlnGly	180
QY	754	TTTACGCCAACCAACCTTTAAGGCCCAATCTAGTGGAAACTAGTTGTCCCAACATCGGATG	813
Db	181	PheThrGlnGlnProLeuAsgProAenLeuValGluThrSerCysProAenIleArgMet	200
QY	814	GATCCAAAATATTATGCCCGCTGATCCAGACTGGATTGCTTTTATATCATGACCAACGATAT	873
Db	201	AspProLysLeuCysProAlaAepProAepTrpIleAlaPheIleHisSerAenAspIle	220
QY	874	TGGATATCTAACATCTGAACACAGAGAGAGAGACTCACATTATGTGCACAACTAGCTA	933
Db	221	TrpIleSerAenIleValThrArgGluLysArgLysLeuTyrValHisGlnGluLeu	240
QY	934	GCCAACTAGGAGAGATGCCAGATCAGCTGGAGTCGCTACCTTGTGTTCTCCCAAGAGAA	993
Db	241	AlaAenMetGluLysAepAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu	260
QY	994	TTTGATAGATATTCGCTATTTGGTGGTGTCCAAAAGCTGAAACAACTCCCAAGTGGTGT	1053
Db	261	PheAepArgTyrSerGlyTyrTrpTyCysProLysAlaGluThrThrProSerGlyGly	280
QY	1054	AAAAATCTTAGAATTTCTATATGAAGAAAATGATGAATCTGAGGTGGAATTTATCATGTT	1113
Db	281	LysIleLeuArgIleLeuTyrGluGluAenAepLysSerGluValGluIleIleHisVal	300
QY	1114	ACATCCCTCATGTTTGGAAAACAGAGGGCCAGATTTCATTCGGTTATCCTTAAAAACAGGTACA	1173
Db	301	ThrSerProMetLeuGluThrArgArgAlaAepSerPheArgTyrProLysThrGlyThr	320
QY	1174	GCAATCCCTAAAGTCATTTTAAGATGTCAGAAAATATGATGATGCTCAAGGAAGGATC	1233
Db	321	AlaAenProLysValThrPheLysMetSerGluIleMetIleAepAlaGluGlyArgIle	340
QY	1234	ATAGATGTCATAGATAAGAACTAATTCACACCTTTTGGATTTCTATTCAAGGAGTTGAA	1293
Db	341	IleAepValIleAepLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu	360
QY	1294	TATATTGCCAGCTGGATGGACTCTCTGAGGAAAATATCTTGGTCCATCTCTACTAGAT	1353
Db	361	TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAep	380

Db 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
Qy 2494 GCTGGGGCCAGTCACCTCTGTGGATCTTATGATACAGGATACACGGAAGCTTATATG 2553
Db 761 AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780
Qy 2554 GGTCAACCTGACAGAAAGAGGGCTATTACTATAGGATCTGTGGCCATGCAAGCAGAA 2613
Db 781 GlyHisProAspGlnAsnGluGlnGlyTyrLeuGlySerValAlaMetGlnAlaGlu 800
Qy 2614 AAGTTCCCTCTGAACCAAAATCGTTTACTCTCTTACATGTTTCTGGATGAGAATGTC 2673
Db 801 LysPheProSerGluProAsnArgLeuLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
Qy 2674 CATTTTGCATACACAGTATATTAAGTGTCTTCTAGTGTGGAGGCTGGAAAGCCATATGAT 2733
Db 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
Qy 2734 TTACAGATCTATCTCAGGAGACACAGCATTAAGAGTTCTCTGAATCGGAGAACATTAT 2793
Db 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
Qy 2794 GAACCTGCATCTTTTGCATCTACCTTCAAGAAAACCTTGGATCAGTATTCCTCTTAAA 2853
Db 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLys 880
Qy 2854 GTGATA 2859
Db 881 ValIle 882

RESULT 2

US-10-070-464-1
; Sequence 1, Application US/10070464
; Patent No. 6881564
; GENERAL INFORMATION:
; APPLICANT: ABBOTT, Catherine Anne
; APPLICANT: CORRELL, Mark Douglas
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES
; FILE REFERENCE: GH-007
; CURRENT APPLICATION NUMBER: US/10/070,464
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: PCT/AU00/01085
; PRIOR FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: AU P05709
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: AU P02762
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-070-464-1

Alignment Scores:
Pred. No.: 0 Length: 882
Score: 4700.00 Matches: 882
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 84.7% Indels: 0
DB: 2 Gaps: 0

US-10-825-632-2 (1-3120) x US-10-070-464-1 (1-882)

Qy 214 ATGGCAGCAGCAATCGAAGACAGACAGCTGGGTGTCAGATATTTGAACCTGGGACTGT 273
Db 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
Qy 274 GAGGAGAATATTGAATCACAGATCGGCCTAAATTTGGAGCTTTTATGCTGAGCGGTAT 333
Db 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40

Qy 334 TCCTGGAGTCAGCTTAAAAAGCTGCTCCGATACCAGAAAATATCATGCTACATCATG 393
Db 41 SerTrpSerGlnLeuLysLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
Qy 394 GCTAAGGACACCATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGACCTCATCA 453
Db 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
Qy 454 GACAGAAATCTATTACCTGCTCCATGTCGTGAGAACACAGAGAAAATACATGTTTATCT 513
Db 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
Qy 514 GAAATTCCTCAAAATCATCAATAGACAGCAGTCTTAATGCTCTCTTGGAGCCCTTTTGT 573
Db 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
Qy 574 GATCTTTTTCAGCAACACTGAGACTATGGAATGTATTCTCGAGAGAGAACTATTAAAGA 633
Db 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140
Qy 634 GAAAGAAAACGCATTTGGAACAGTCGGAATTGCTTCTTACGATTATCACCAAGGAAGTGA 693
Db 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
Qy 694 ACATTTCTGTTTCAAGCCGTTAGTGAATTTATACGTAAAGATGAGAGGCCCAAGGA 753
Db 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
Qy 754 TTTAGCAACAACCTTTTAAAGGCCCAATCTAGTGGAAACTAGTTTCTCCACATACGGATG 813
Db 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
Qy 814 GATCCAAAATATGCCCCGCTGATCCAGACTGGATTGCTTTTATACATAGCAACGATATT 873
Db 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAspAspIle 220
Qy 874 TGGATATCTAACATCGTTAACAGAGAGAGAGAGACTCCTTATGTGACAAATGAGCTA 933
Db 221 TrpIleSerAsnIleValThrArgGluArgLeuThrTyrValHisAsnGluLeu 240
Qy 934 GCCAATCGAAGAAAGATGCCAGATCAGCTGAGTGCCTACCTTTGTTTCTCAAGAGAA 993
Db 241 AlaAsnMetGluGluLysAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu 260
Qy 994 TTTGATAGATATTCGGCTATTGGTGTCTCAAAAGCTGAGAAACAACTCCAGTGGTGGT 1053
Db 261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280
Qy 1054 AAAATTTCTAGAAATCTATATGAAGAAAATGATGATCTGAGGTGGGAAAATTTATTCATGTT 1113
Db 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHisVal 300
Qy 1114 ACATCCCCTATGTTGGAAAACAGGAGGCGAGATTCATTCCCTTATCTTAAACAGGTACA 1173
Db 301 ThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThr 320
Qy 1174 GCAATCTTAAAGTCATTTTAAAGTGTCAAGAAATGATGATGCTGAGGAGGAGATC 1233
Db 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
Qy 1234 ATAGATGTCATAGATAAGGAACCTTAATCAACCTTTTGAGATTCCTTATTTGAAGGAGTTGAA 1293
Db 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
Qy 1294 TATATTGCCAGAGCTGGATGGACTCCTCGAGGAAAATATGCTTGGTCCATCTACTAGAT 1353
Db 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
Qy 1354 CGCTCCAGACTCGCCTACAGATGTTGTGATCTCACCTGAAATTTATTTATCCCAAGTGA 1413
Db 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400

QY	1414	GATGATGTTATGAAAGGACGAGACTCATTGAGTCAGTGCCTGATTCCTGTGACGCCACTA	1473	Db	761	AlaGlyAlaProValThrLeuThrPheThrAspThrGlyThrGluArgTyrMet	780
Db	401	AspAspValMetGluArgGlnArgLeuGluLeuGluSerValProAspSerValThrProLeu	420	QY	2554	GGTCACCCCTGACCAAGATGAACAGGGCTATTACTTAGGATCTCTGCCCATGCAAGCAGAA	2613
QY	1474	ATTATCTATCAAGAAACAACAGACATCTCGGATAAATATCCATGACATCTTTCATGTGTTTT	1533	Db	781	GlyHisProAspGlnAenGluGlnGlyTyrLeuGlySerValAlaMetGlnAlaGlu	800
Db	421	IleIleTyrGluGluThrThrAspIleTyrPheAenIleHisAspIlePheHisValPhe	440	QY	2614	AAGTTCCCTCTCGAACCAAAATCGTTTACTGCTCTTACATGGTTTCTCTGATGAGAATGTC	2673
QY	1534	CCCCAAGTCACGAAGGAAATTTGAGTTTATTTTGGCTCTGATGCAAAACAGGTTTC	1593	Db	801	LysPheProSerGluProAenArgLeuLeuLeuHisGlyPheLeuAspGluAenVal	820
Db	441	ProGlnSerHisGluGluGluIleGluPheIlePheAseSerGluCysAlaThrGlyPhe	460	QY	2674	CATTTTGCACATACCAGTATATTAAGTCTTCTAGTGGGCTGGAAGCCATATGAT	2733
QY	1594	CGTCATTTATACAAATATACATCTATTTAAAGGAAGCAATATAAAGCATCCAGTGGT	1653	Db	821	HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp	840
Db	461	ArgHisLeuTyrIleThrSerIleLeuLeuGlySerLysTyrLysArgSerSerGly	480	QY	2734	TTACAGATCTATCTCTCAGGAGACACAGACATGAAGTTCCTGAATCGGAGAACATTAT	2793
QY	1654	GGGTGCTGCTCCAGTGAATTTCAAGTGTCTTATCAAGAGGATAGCAATACCAAGT	1713	Db	841	LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr	860
Db	481	GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer	500	QY	2794	GAACTGCATCTTTTGCATCTTCAAGAAAAACCTTGGATCAGGTATTTGCTCTTAAAA	2853
QY	1714	GGTGAATGGGAAGTTCTTGCCGGCATGGATCTAATATCAAGTTGATGAAGTCGAAGG	1773	Db	861	GluLeuHisLeuLeuHisTyrLeuGlnGluAenLeuGlySerArgIleAlaAlaLeuLys	880
Db	501	GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg	520	QY	2854	GTGATA 2859	
QY	1774	CTGTTATATTTTGAAGCCACAAAGACTCCCTTTTAGAGCATCACTGTACGTAGTCAGT	1833	Db	881	ValIle 882	
Db	521	LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSer	540	RESULT 3			
QY	1834	TACGTAAATCTCGAGAGGTGACAAAGGCTGACTGACCGTGGCTACTCACATTTCTGTGC	1893	US-09-976-674-7			
Db	541	TyrValAenProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys	560	; Sequence 7, Application US/09976674			
QY	1894	ATCAGTCAGCACTGTGACTTCTTTATAGTAAGTATAGTAAGCAAGAAATCCACACTGT	1953	; Patent No. 6844180			
Db	561	IleSerGlnHisCysAspPhePheIleSerLysTyrSerAenGlnLysAenProHisCys	580	; GENERAL INFORMATION:			
QY	1954	GTGTCCTTTACAGCTATCAAGTCTCGAAGATGACCAACTTGCAAAACAAAGAAATTT	2013	; APPLICANT: Qi, Steve			
Db	581	ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe	600	; APPLICANT: Akinsanya, Karen			
QY	2014	TGGGCCACATTTTGGATTCAGCAGGTCTCTTCTGACTATATCTCTCCAGAAATTTTC	2073	; APPLICANT: Riviere, Pierre			
Db	601	TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe	620	; APPLICANT: Junien, Jean-Louis			
QY	2074	TCTTTTGAAGTACTGATGATTTTACATTTGATTTGGATGCTCTACAGCCTCATGCTA	2133	; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV			
Db	621	SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu	640	; FILE REFERENCE: 70669			
QY	2134	CAGCTGGAAGAAATATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2193	; CURRENT APPLICATION NUMBER: US/09/976,674			
Db	641	GlnProGlyLysLysTyrProThrValIlePheIleTyrGlyGlyProGlnValGlnLeu	660	; PRIOR FILING DATE: 2001-10-12			
QY	2194	GTGAATAATCCGTTTAAAGGAGTCAAGTATTTCGCTTGAATACCCCTAGCCTCTAGGT	2253	; PRIOR APPLICATION NUMBER: US 60/240,117			
Db	661	ValAenAenArgPheLysGlyValLysTyrPheArgLeuAenThrLeuAlaSerLeuGly	680	; PRIOR FILING DATE: 2000-10-12			
QY	2254	TATGTGTTGTAGTGATAGACAAACAGGGGATCTGTGTCACCGAGGCTTAAATTTGAAGC	2313	; NUMBER OF SEQ ID NOS: 61			
Db	681	TyrValValValValIleAenAenArgLysSerCysHisArgGlyLeuLysPheGluGly	700	; SOFTWARE: PatentIn version 3.1			
QY	2314	GCCTTTAAATATAAAATGGGTCAATAGAAATTCAGCATCAGGTGGAAGACTTCCAAAT	2373	; SEQ ID NO 7			
Db	701	AlaPheLysTyrLysMetGlyGlnIleGluIleAenAspGlnValGluGlyLeuGlnTyr	720	; LENGTH: 690			
QY	2374	CTAGCTCTCGATATGATTTTCAATTGATCTAGATCGTGTGGCATCCACGGCTGCTCTAT	2433	; TYPE: PRT			
Db	721	LeuAlaSerArgTyrAspPheIleAenLeuAspArgValGlyIleHisGlyTyrPheSerTyr	740	; ORGANISM: Homo sapiens			
QY	2434	GGAGGATACCTCTCCCTGATGATTAATGACAGAGTTCAGATATCTTCAGGGTTGCTATT	2493	US-09-976-674-7			
Db	741	GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle	760	Alignment Scores:			
QY	2494	GCTGGGGCCCCAGTCACCTGTGGATCTTCTATGATACAGGATACACGGAACGTTATATG	2553	Pred. No.: 0			
				Score: 3607.50			
				Percent Similarity: 87.1%			
				Best Local Similarity: 87.1%			
				Query Match: 65.0%			
				DB: 2			
				US-10-825-632-2 (1-3120) x US-09-976-674-7 (1-690)			
				QY 214 ATGGCAGCAGCAATGGAACACAGACGCTGGTGTGAGATATTTGAACTCGGACTGT			
				Db 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys			
				QY 274 GAGGAGAATTTGAATCACAGGATCGCCCTAAATTTGGAGCCCTTTTATGTTGAGGGTAT			
				Db 21 GluGluAenIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr			
				QY 334 TCCTGGAGTCAGCTTAAAAAGCTGTGTCGGATACACAGAAATATCATGGCTACATGATG			
				Db 41 SerTrpSerGlnLeuLysLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet			
				QY 394 GCTAAGGCACCATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGGAGCTCATTC			

Db 61 AlaLysAlaProHisAspPheMetPheValIysArgAsnAspProAspGlyProHisSer 80
QY 454 GACAGAACTATTACTCTGGCCATGTCGTGGTGAACACAGAGAAAATACACTGTTTATTCT 513
Db 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
QY 514 GAAATTTCCAAACTATCAATAGACGACAGTCTTAATGCTCTCTGGAGCCCTCTTTTG 573
Db 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
QY 574 GATCTTTTTCAGCAACACTGCACTATGGAATGTATTCTCGAGAAGAAGAACTATTAAAG 633
Db 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140
QY 634 GAAAGAAAACGCAATTTGGAACAGTCGGAATTCCTTTACGATTATCACCAAGGAAGTGA 693
Db 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
QY 694 ACATTTCGTGTTTCAAGCCGGTAGTGAATTTATCACGTAAAGATGAGAGGCCACAAAG 753
Db 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
QY 754 TTTAGCCAAACCTTTTAAGCCCMATCTAGTGAAACTAGTTGTCCCAACATACCGATG 813
Db 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
QY 814 GATCCAAATATGATCCCGCTGATCCAGACTGGATTGCTTTTATACATAGCAACGATATT 873
Db 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
QY 874 TCGATATCTAACATCGTAACACAGAGAAAGAGACTCACTTATGTGCAACATAGCTA 933
Db 221 TrpIleSerAsnIleValThrArgGluGluArgGluLeuThrTyrValHisAsnGluLeu 240
QY 934 GCCAATCGGAAGAAGATGCCAGATCAGCTGAGTGCCTACCTTGTCTCCCAAGAGAA 993
Db 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluLeu 260
QY 994 TTTGATPAGATATTCGGCTATTGTTGGTGTCCAAAGCTGAAACACTCCCGAGTGTGT 1053
Db 261 PheAspArgTyrSerGlyTyrTrpCysProLysAlaGluThrThrProSerGlyGly 280
QY 1054 AAAATTTAGAAATCTATATGAAGAAAATGATGAATCTGAGGTGGAAATTTATCATGTT 1113
Db 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300
QY 1114 ACATCCCTATGTTGGAACAAGAGGCGCAGATTCAATCCGTATCCGTAAACAGGTACA 1173
Db 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320
QY 1174 GCAAATCTTAAGTCACTTTTAAGATGTCAGAAATAATGATTGATGCTGAAGGAAGATC 1233
Db 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
QY 1234 ATAGATGTCATAGATAAGAACTAATTCACCTTTTGAGATTCATTGGAAGAGTTGAA 1293
Db 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
QY 1294 TATATTGCCAGCTGGATGGACTCCTGAGGGAAAATATGCTTGGTCCATCCTACTAGAT 1353
Db 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
QY 1354 CGTCCCGACACTCGCTACAGATAGTGTGATCTCACCTGAAATTTATTTATCCAGTAGAA 1413
Db 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
QY 1414 GATGATGTTATGGAAGGACAGACTCAATTGATGACGTGCTGATTCGTGACGCCACTA 1473
Db 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
QY 1474 ATTATCTATGAAGAAACAACACATCTGGATAATATATCCATGATCTTTTCATGTTTT 1533

Db 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
QY 1534 CCCCAAGTCACGAAGCAGAAATTGAGTTATTTTGGTCTCTGAATGCAAAACAGGTTTC 1593
Db 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
QY 1594 CGTCATTATACAAAATTACATCTATTTTAAAGGAAAGCAATATATAACGATCCAGTGGT 1653
Db 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
QY 1654 GGGCTGCTGCTCCAAAGTGAATTTCAAGTGTCTTATCAAGAGGAGATGACGAATACCACT 1713
Db 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 500
QY 1714 GGTGAATGGGAAGTTCTTGGCCGCATGGAATCTAATATCCAAGTTGATGAAGTCAGAAG 1773
Db 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
QY 1774 CTGGTATATTTGAAGGCCAACAAAGACTCCCTTTTAGAGCATCACCTGTACGTACTAGT 1833
Db 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisIleLeuTyrValValSer 540
QY 1834 TACCTAAATCCTCGACAGGTGCAAGCTGACTGACCGTGGCTACTCACATTTCTGTGTC 1893
Db 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
QY 1894 ATCAGTCAGCACTGTGACTTCTTTATAGTAAGTATAGTAACACAGAAGAATCCACACTGT 1953
Db 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
QY 1954 GTGTCCCTTTTAAAGCTATCAAGTCTGAAGATGACCCAACTTGGCAAAACAAAGAAATTT 2013
Db 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
QY 2014 TGGGCCACCAATTTGGATTACAGAGTCTCTCTCTCTGACTATATCTCTCCAGAAATTTTC 2073
Db 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 620
QY 2074 TCTTTTGAAGTACTACTGGATTTACATTTGATGGATGCTCTACAAGCCTCATGATCTA 2133
Db 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
QY 2134 CAGCTCGAAGAAATATCTACTGTGTTTCAATATATGTTGGTCTCTCAGGTGCAAGTTG 2193
Db 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGly 655
QY 2194 GTGAATAACGGTTTAAAGGAGTCAAGTATTCCCGCTTGAATACCTCTCTAGCTCTCTAGT 2253
Db 655 655
QY 2254 TATGTGTTGTAGTGATAGACAAACAGGGGATCCTGTACCGAGGGCTTAAATTTGAAGC 2313
Db 655 655
QY 2314 GCCTTTAAATATATAAATGGGTCAATAGAAAATTGACGATCAGGTGGAAGGACTCCAATAT 2373
Db 655 655
QY 2374 CTAGCTTCGATATGATTTTCATTGACTTAGATCGTGTGGGCATCCACGGCTGGTCTAT 2433
Db 655 655
QY 2434 GGAGGATACCTCTCCCTGATGGCATTAATGACAGAGTGCAGATATCTTCAGGGTGTGCTATT 2493
Db 656 659
QY 2494 GCTGGGGCCCGAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAACGTTATATG 2553
Db 659 uLeuGlyProGlnSerLeuCysGlySerSerMetIleGlnAspThrArgAsnValIleTr 679
QY 2554 GGTCAACCTGACAGAAATGAACAGGCTATTACT 2587
Db 679 pValThrLeuThrArgMetAsnArgAlaIleThr 690

Tue Apr 18 08:18:35 2006

```
RESULT 4
US-09-976-674-11
; Sequence 11, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976, 674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 11
; LENGTH: 661
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-11

Alignment Scores:
Pred. No.: 0 Length: 661
Score: 3513.50 Matches: 660
Percent Similarity: 99.7% Conservative: 0
Best Local Similarity: 99.7% Mismatches: 0
Query Match: 63.3% Indels: 2
DB: 2 Gaps: 1

US-10-825-632-2 (1-3120) x US-09-976-674-11 (1-661)
QY 214 ATGGCAGCAGCAATGGAACAGACAGCTGGTGTGAGATATTTGAACTCGGACTGT 273
Db 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
QY 274 GAGGAGATATGATCAGAGTCGCGCTAAATTCGAGCGCTTTTATCTTCAGCGGTAT 333
Db 21 GluGluIleGlnSerGlnAspArgProIysLeuGluProPheTyrValGluArgTyr 40
QY 334 TCCTGGAGTCAGCTAAAGCTCTCCGATACCGAAGAAATATCATCGCTACATGATG 393
Db 41 SerTrpSerGlnLeuLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
QY 394 GCTAAGGCCACATGATTTTCATGTTGTGAAGGAATGATCCAGATGGACTCATTTCA 453
Db 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
QY 454 GACAGAATCTATTACCTTGCCTGCTGCTGAGAACAGAGAAATACACTGTTTATTCT 513
Db 81 AspA-glieTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
QY 514 GAAATTCACAAACTATCAATAGACGACAGCTCTTAATGCTCTCTTGGAGCGCTCTTTTG 573
Db 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
QY 574 GATCTTTTTCAGCAACACTCGACTATGGAATGATTTCTCGAGAAGAAGAACTATTAAAG 633
Db 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg 140
QY 634 GAAGAAACGATTTGGACAGTCGGATTCCTTACGATTATCACCAAGGAAGTGA 693
Db 141 GluA-glyArgIleGlyThrValGlyIleAspTyrAspTyrHisGlnGlnSerGly 160
QY 694 ACNTTCTGTTTCAGCGGTAGGTGANTTTATCACTTAAAGATGAGGCGCACAGGA 753
Db 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
QY 754 TTTACGCAACACTTTAAGCCCATCTAGTGAAGAACTAGTTGTCACACATACCGATG 813
Db 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
814 GATCCAAATTTATGCCCGCTGATCCAGACTGGATTCCTTTTATACATACAGCATATTT 873
Db 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
QY 874 TGGATATCTAAACATCGTTAACCCAGAGAAGAGAGACTCACTTATGTGCAATGAGCTA 933
Db 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu 240
QY 934 GCCAACATCGGAAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTCTCCAAAGAA 993
Db 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu 260
QY 994 TTTTGATAGATATTTCTGGCTATTGTGTGTCCTCAAAAGCTGAAACAACTCCAGTG 1053
Db 261 PheAspArgTyrSerGlyTyrTrpCysProLysAlaGluThrThrProSerGlyGly 280
QY 1054 AAAAACTTTAGAAATCTATATGAAGAAAAATGAAATCTGAGGTGGAAAAATTTATCATG 1113
Db 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300
QY 1114 ACATCCCTATGTTGGAAACAGAGGCGCAGATTTCATTCGTTTATCTTAAACAGGTACA 1173
Db 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320
QY 1174 GCAATCTCTAAAGTCACCTTTTAAGATGTCAGAAATAATGATGATGCTGAAGGAAGATC 1233
Db 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
QY 1234 ATAGATGTCATAGATAAGGAACCTAATCAACCTTTGAGATCTCTATTGAGGAGTTGAA 1293
Db 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
QY 1294 TATATTGCCAGAGCTGGATGGAATCCTCGAGGAAAAATATGCTGCTCCATCTACTAGAT 1353
Db 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
QY 1354 CGCTCCACAGACTCGCTACAGATAGTTGATCTCACCTGAATTTATTTATCCCAAGTAA 1413
Db 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
QY 1414 GATGATGTTATGGAAGGAGAGACTCATGAGTCAGTCGCTGATTTCTGTGACGCCACTA 1473
Db 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
QY 1474 ATTATCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTT 1533
Db 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
QY 1534 CCCAAAGTCAGAGAGGAAATTCAGTTTATTTTGGCTCTGAATGCAAAACAGGTTTC 1593
Db 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
QY 1594 CGTCATTTATACAAATTCATCTATTTTAAAGGAAGCAATATTAACGATCCAGTGGT 1653
Db 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
QY 1654 GGGCTGCTGCTCCAAAGTGAATTCAGTGTCTTATCAAAGAGAGATAGCAATACCAGT 1713
Db 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleHisSer 500
QY 1714 GGTGAATGGGAAGTCTTCTGCGCGCATGGATCTAATATCCAAAGTTGATGAAGTCGAAG 1773
Db 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
QY 1774 CTGATATATTTTGAAGGCCCAAGACTCCCTTTTAGAGCATCACCTGTACCTAGTCACT 1833
Db 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSer 540
QY 1834 TACGTAATCTCTGAGAGGTGACAAGCTGACCTGACCGTGGCTACTCATCTTCTGCTGC 1893
Db 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCys 560
```

1894 ATCAGTCAGCAGTGTGACTTCTTTATAGTAAGTATAGTAACCAAGAAATCCACACTGT 1953
Db |||||
561 IleserGlnHisCysaspPhePheHisSerLysTyrSerAsnGlnLysAsnProHisCys 580
Qy |||||
1954 GTGTCCCTTTTACCAAGCTATCAAGTCTCTGAAGATGACCCCAACTGTGCMAAACCAAGAAATTT 2013
Db |||||
581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
Qy |||||
2014 TGGGCCACCAATTTGGATTGAGAGTCTCTCTCTCTGCTATGCTATGCTCTCCAGAAATTTTC 2073
Db |||||
601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 620
Qy |||||
2074 TCTTTTCAAGTACTGATTTACATTTGATGGAGTCTCTCAAGCCCTCATGACTA 2133
Db |||||
621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
Qy |||||
2134 CAGCTGGAAGAAATATCTACTGCTGCTTCATATATGGTGGTCTC---AGTGGCAG 2190
Db |||||
641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyLeuLeuArgCysSe 660
Qy |||||
2191 TTGG 2194
Db |||||
660 ITrp 661
RESULT 5
; Sequence 19, Application US/09976674
; Patent No. 6944180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
; LENGTH: 658
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-19
Alignment Scores:
Pred. No.: 0 Length: 658
Score: 3504.00 Matches: 655
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 63.1% Indels: 0
DB: 2 Gaps: 0
US-10-825-632-2 (1-3120) x US-09-976-674-19 (1-658)
Qy 214 ATGGCAGCAGCAATGGAAACAGACAGCTGGGTGTGAGATATTGAACTGGGACTGT 273
Db 1 MetaAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
Qy 274 GAGGAGATATTGATACAGATCGGCTAAATTTGGAGCTTTTATGTTGAGCGGTAT 333
Db 21 GluGluAenIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
Qy 334 TCCTGGAGTCAGCTTAAAGCTGCTTGGCCGATACCAAGAAATATCATGCTCATGATG 393
Db 41 SerTrpSerGlnLeuLysLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
Qy 394 GCTAAGGCCACCATGATTGTTGTGTAAGAGGAATGATCCAGATGACCTCATTTCA 453
Db |||||
61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80

Qy 454 GACAGAAATCTATTACCTTGGCCATGCTCGGTGAGACAGAGAAAATACACTGTTTATTCT 513
Db |||||
81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
Qy |||||
514 GAAATTCCTCAAAATCATCAATAGACAGCAGTCTTAATGCTCTCTTGGAGCCCTCTTTTG 573
Db |||||
101 GluIleProLysThrIleAsnArgAlaValLeuMetLeuSerTrpLysProLeuLeu 120
Qy |||||
574 GATCTTTTTCGAGCAACACTGGAATGATTTCTCGAGAAGAAGAACTATTAAAGA 633
Db |||||
121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140
Qy |||||
634 GAAAGAAAACGATTTGGAACAGTCCGGAATTCCTTTACGATTATACCAAGGAAGTGA 693
Db |||||
141 GluArgLysArgIleGlyThrValGlyIleAserTyrAspTyrHisGlnGlySerGly 160
Qy |||||
694 ACATTTCTGTTTCAAGCGGTAGTGGAAATTTATCACGTAAAGATGAGAGGCCCAAGGA 753
Db |||||
161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
Qy |||||
754 TTTAGCAACAACCTTTAAGGCCCAATCTAGTGGAACTAGTCTCCCAACATACGCGATG 813
Db |||||
181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
Qy |||||
814 GATCCAAAATTTATGCCCGCTGATCCAGACTGGATTGCTTTTATACATAGCAACGATATT 873
Db |||||
201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
Qy |||||
874 TGGATATCTAACATCGTTAAACCAAGAGAAGAAAGAGACTCACTTATGTGCAATGAGCTA 933
Db |||||
221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisGlnGluLeu 240
Qy |||||
934 GCCAATGGAAGAAGATGCCAGATCAGCTGAGTCCGCTACCTTTGTTCTCCAAAGAA 993
Db |||||
241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluLeu 260
Qy |||||
994 TTTGATPAGATATCTGGCTATTGGTGTGTCGAAAGCTGAAACAACTCCAGTGGTGT 1053
Db |||||
261 PheAspArgTyrSerGlyTyrTrpCysProLysAlaGluThrThrProSerGlyGly 280
Qy |||||
1054 AAAATCTTAGAATCTATGAAGAAATGATGAACTGAGGTGGAAATTTATTCATGTT 1113
Db |||||
281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHisVal 300
Qy |||||
1114 ACATCCCTTATGTTGAAACAAGGAGGCGAGATTCATTCCGTATCTCTTAAACACAGTACA 1173
Db |||||
301 ThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThr 320
Qy |||||
1174 GCAATCTCTAAAGTCACCTTTTAAAGATGTCAGAAATTAATGATTGATGCTGAAGAGGATC 1233
Db |||||
321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
Qy |||||
1234 ATAGATGTCATAGATAAGGAACCTAAATCAACCTTTTGAGATTCATTTTGAAGAGTGGAA 1293
Db |||||
341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
Qy |||||
1294 TATATTGCCAGACTGGATGGAATCTCTGAGGAAAAATATGCTGTGCTCCATCTACTAGAT 1353
Db |||||
361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
Qy |||||
1354 CGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAAATTTATTTATCCAGTAGAA 1413
Db |||||
381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
Qy |||||
1414 GATGATGTTATGAAAGGAGCAGACTCATTTGAGTCAGTCCCTGATTTCTGTGACGCCACTA 1473
Db |||||
401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
Qy |||||
1474 ATTATCTATCAAGAAAAACACAGACATCTCGATAAATATCCATGACATCTTTTCATGTTT 1533
Db |||||
421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
Qy |||||
1534 CCCAAAGTCACCAAGAGGAAATTTAGTTATTTTTCCTCTGAAATGCAAAACAGGTTC 1593

|||||
Db 441 ProGlnSerHisGluGluGluGluGluPhePheAlaSerGluCysLysThrGlyPhe 460
QY 1594 CGTCATTATACAAATATACATATTTTAAAGGAAGCAATATATAACGATCCAGTGCT 1653
Db 461 ArgHisLeuTyrLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 480
QY 1654 GGGCTGCTCTCCAGTGAATTTCAAGTGTCTCTATCAAGAGGAGATAGCAATACCAAGT 1713
Db 481 GlyLeuProAlaProSerAspPheLysCysProLysGluGluGluGluGluGluGluGlu 500
QY 1714 GGTGAATGGGAAGTCTTGGCCGGCATGGATCTAATATCAAGTGTGAAGTCAAGAGG 1773
Db 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnLysGluValAspGluValArgArg 520
QY 1774 CTGTATATTTTGAAGCACCAAGACTCCCTTTAGAGCATCACTGTACGTAGTCAAGT 1833
Db 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer 540
QY 1834 TACGTAAATCTCGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACATTTCTTGCTGC 1893
Db 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
QY 1894 ATCAGTCAGCACTGTGACTTCTTTATAAGTAAAGTATAGTAACCAAGAGAAATCCACACTGT 1953
Db 561 IleSerGlnHisCysAspPhePheLysSerLysTyrSerAsnGlnLysAsnProHisCys 580
QY 1954 GTGTCCTTTTACAAAGCTATCAAGTCTGAAGATGACCCCAACTTGCAAAACAAAGGAATTT 2013
Db 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
QY 2014 TGGGCCACCACTTTTGGATTGAGTTCAGCAGGTCTCTCTGACTATACCTCCAGAAATTTTC 2073
Db 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluLysPhe 620
QY 2074 TCITTTGAAAGTACTGATTTACATTTGATGGATGCTCTACAGCCCTCATGATCTA 2133
Db 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
QY 2134 CAGCTCGAAGAAATCTCTACTGTGCTGTTCATATATGFGGT 2178
Db 641 GlnProGlyLysLysTyrProThrValLeuPheLysIleTyrGlyGly 655

RESULT 6
; Sequence 21, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 21
; LENGTH: 613
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-21

Alignment Scores:
Pred. No.: 0 Length: 613
Score: 3236.00 Matches: 607
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 58.3% Indels: 0
DB: 2 Gaps: 0

US-10-825-632-2 (1-3120) x US-09-976-674-21 (1-613)
QY 214 ATGGCAGCAGCAATGGAACACAGACACTGGCTGTCAGATATTTGAACTCGGACTGT 273
Db 1 MetAlaAlaAlaMetGluThrGluGluLeuGlyValGluLysPheGluThrAlaAspCys 20
QY 274 GAGGAGATATTTGAATCAAGATCGGCTTAATTTGGAGCCTTTTATGTTGAGCGGTAT 333
Db 21 GluGluAsnLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 40
QY 334 TCTCGAGTACAGCTTAAAGACCTGCTTCGATACCAAGAAATATCATGCTCATGATG 393
Db 41 SerTrpSerGlnLeuLysLysLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
QY 394 GCTAAGGCAACCATGATTTTCAATGTTGTGAAGAGGAATGATCCAGATGAGCCTCATPCA 453
Db 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
QY 454 GACAGAATCTATTACCTTTGCCATCTGCTGGTGAGAACACAGAGAAATACACTGTTTATTCT 513
Db 81 AspArgLysTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
QY 514 GAAATTTCCAAATCTATCAATAGACGACAGTCTTAATGCTCTCTTGGAAAGCCTCTTTTG 573
Db 101 GluLysProLysThrIleAsnArgAlaValLeuMetLeuSerTrpLysProLeuLeu 120
QY 574 GATCTTTTTCAGGCAACACTCGACTATGGAATGTATTCTCGAAGAGAACTATTAAAGA 633
Db 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140
QY 634 GAAAGAAACCATTTGGAACAGTCCGAAATGCTTCTTACGATATATCACCAAGGAAGTGA 693
Db 141 GluArgLysArgLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 160
QY 694 ACATTTCTGTTTCAAGCGGTAGTCAAGTATTTATCAGTAAAGTGAAGGAGGACACAGGA 753
Db 161 ThrPheLeuPheGlnAlaGlySerGlyLysLysLysLysLysLysLysLysLysLys 180
QY 754 TTTACGCAACACCTTTAAGCCCAATCTAGTGAAACTAGTGTGTCCTCCACATACCGATG 813
Db 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnLysLeuArgMet 200
QY 814 GATCCAAATATGATGCTCCGCTGATCCAGACTGGAGTCTTTTATACATACAAACGATATT 873
Db 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
QY 874 TGGATATCTAAATCGTAAACCAAGAGAGAGACTCACTTATGTGCAATAGAGCTA 933
Db 221 TrpIleSerAsnLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 240
QY 934 GCCAACATGGAAGAGATGCCAGATCAGCTGGAGTCCCTACCTTGTCTCCCAAGAGAA 993
Db 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGluGlu 260
QY 994 TTTGATAGATATTTCTGGCTATTGCTGTTGTCCTCCAAAGCTCAACCTCCAGTGGTGT 1053
Db 261 PheAspArgTyrSerGlyTyrTrpCysProLysAlaGluThrThrProSerGlyGly 280
QY 1054 AAAATTTCTAGAAATTTCTATATGAAGAAATGATGAATCTCAGGTGGAAATTTATTCATGTT 1113
Db 281 LysIleLeuArgLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 300
QY 1114 ACATCCCTATGTTGGAACCAAGAGGAGGAGATTCATTCCGCTTATCTTAAACAGAGTACA 1173
Db 301 ThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThr 320
QY 1174 GCAAATCTTAAGTCACTTTTAAAGTGTCAAGAAATATGATGCTGAGGAAGGATC 1233
Db 321 AlaAsnProLysValThrPheLysMetSerGluLysMetIleAspAlaGluGlyArgLys 340
QY 1234 ATAGATGTCATAGATAAGGAATCAATTAACCTTTTGGATTTCTATTGAGGAGTTGAA 1293

Db 341 IleaspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
Qy 1294 TATATTGCCAGAGCTGGATGAGTCTCTGAGGAAAAATATGCTTGGTCCATCTACTATGAT 1353
Db 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
Qy 1354 CGCTCCAGAGTCCGCTCAGACAGTAGTGGTATCTCAGCTGAATATTTATCCAGTAGAA 1413
Db 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
Qy 1414 GATGATCTTATGAAGGAGAGACATCTAGTCTAGTCTGCTCTGATCTTCTGTGACGCCACTA 1473
Db 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
Qy 1474 ATTATCTATCAAGAAACAAACAGACATCTGGATAAATATCCATGACATCTTTCATGTTTTT 1533
Db 421 IleIleTyrGluGluThrThrAspIleTyrIleAsnIleHisAspIlePheHisValPhe 440
Qy 1534 CCCAAAGTCACGAAGAGAAATGTAGTTTATTTTGGCTCTGGAATGCAAAACAGGTTTC 1593
Db 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
Qy 1594 CGTCATTTATCAAAATATCATCTATTTTAAAGGAAGCAAAATATAAAGCATCCAGTGGT 1653
Db 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
Qy 1654 GGGCTGCTCTCCAGTGTATTTCAAGTGTCTTATCAAGAGAGAGATAGCAATTACCAGT 1713
Db 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 500
Qy 1714 GGTCAATGGGAAGTCTTGGCCGCATGATCTAATATCAAGTTGATGAAGTCAGAAAGG 1773
Db 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
Qy 1774 CTGTATATTTTGAAGCACCAAGACTCCCTTTTAGAGCATCACCTGTACGTAGTCACT 1833
Db 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisIleLeuTyrValValSer 540
Qy 1834 TACGTAAATCTCGAGAGGTGACAAAGCTGACCTGACCGGTGGCTACTACATCTTGTGTCG 1893
Db 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
Qy 1894 ATCAGTCAGACTGTGACTCTTTATTAAGTAACTATAGTAAACAGAGAAATCCACACTGT 1953
Db 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
Qy 1954 GTGTCCCTTTACAGCTATCAAGTCTCTGAGATGCCCAACTTGCAAAACAAAGGAATTT 2013
Db 581 ValSerLeuTyrLysLeuSerSerProGluAspProThrCysLysThrLysGluPhe 600
Qy 2014 TGGGCCACCATTTTGGATTCA 2034
Db 601 TrpAlaThrIleLeuAspSer 607

RESULT 7
US-09-976-674-23
; Sequence 23, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: OI, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 23
; LENGTH: 892

; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-23
Alignment Scores: 9.06e-296 Length: 892
Pred. No.: 2871.00 Matches: 529
Score: 73.7% Conservative: 137
Best Local Similarity: 58.5% Mismatches: 208
Query Match: 51.7% Indels: 30
DB: 2 Gaps: 5
US-10-825-632-2 (1-3120) x US-09-976-674-23 (1-892)
Qy 124 TGGAGCGCGGCAGCATGAAGCGGCGAGCGGCTCCATAGCGACGTCGCGACGGTCC 183
Db 17 TrpArgSerPheSerLeuAsnSerGluGlyAlaGluArgMetAlaThrThrGlyThrPro 36
Qy 184 GGGCGGGCGCGGGGGAAGAAATGCAACATGCGCAGCAGCAATGGAAACAGAACAGCTG 243
Db 37 ThrAlaAspArgGlyAsp-----AlaAlaAla----- 45
Qy 244 CGTGTGTGAGATATTTGAAACTGCGGACTGTGAGGAGAATATTGAATCACAGGATCGCGCT 303
Db 46 -----ThrAspAspPro 49
Qy 304 AAATTGGAGCCTTTTATGTTGAGCGGTATTTCTGGAGTCAGCTTAAAAAGCTGCTTGC 363
Db 50 AlaAlaArg---PheGlnValGlnLysHisSerTrpAspGlyLeuArgSerIleIleHis 68
Qy 364 GATACAGAAATATATCATGCTGCTACATGATGCTAAGGCACACCATGATTTTCATGTTGTG 423
Db 69 GlySerArgLysTyrSerGlyLeuIleValAsnLysAlaProHisAspPheGlnPheVal 88
Qy 424 AGAGAGATGATCCAGATGGACCTCATTCACAGAGAATCTATTACCTTGCCTCTCTGGT 483
Db 89 GlnLysThrAspLysSerGlyProHisSerHisArgLeuTyrTyrLeuGlyMetProTyr 108
Qy 484 GAGAACAGAGAAATATCACTGTTTATTTCTGAAATTTCCAAATCTATCAATAGACAGCA 543
Db 109 GlySerArgLysSerLeuLeuTyrSerGluIleProLysLysValArgLysGluAla 128
Qy 544 GTCTTAATCTCTCTTGAAGCCTCTTTTGGATCTTTTTCAGCAACATGCTGCTATGGA 603
Db 129 LeuLeuLeuLeuSerTrpLysGlnMetLeuAspHisPheGlnAlaThrProHisGly 148
Qy 604 ATGATATTCGAGAGAGAACTATTAAGAGAAAGAAACGATTTGGACAGCTCGGAATT 663
Db 149 ValTyrSerArgGluGluLeuLeuArgGluArgLysArgLeuGlyValPheGlyIle 168
Qy 664 GCTTCTTACGATTATCCCAAGAAAGTGAACATTTCTTCAAGCCGGTAGTGAATT 723
Db 169 ThrSerTyrAspPheHisSerGlyLeuPheGluPheGlnAlaSerAsnSerLeu 188
Qy 724 TATCAGTAAAGATGGAGGCCACAGGATTTACGCAACACTTTAAGGCCCAATCTA 783
Db 189 PheHisCysArgAspGlyGlyLysAsnGlyPheMetValSerProMetLysProLeuGlu 208
Qy 784 GTGGAACACTAGTTGTCCCAACATACGATGGATCCAAAATTTATGCCCGCTGATCCAGAC 843
Db 209 IleLysThrGlnCysSerGlyProArgMetAspProLysIleCysProAlaAspProAla 228
Qy 844 TGATTTGCTTTTATACATAGCAACGATATTTGGATATCTTAACATCGTAACAGAGAGAA 903
Db 229 PhePheSerPheIleAsnAsnSerAspLeuTrpValAlaAsnIleGluThrGlyGlu 248
Qy 904 AGGAGACTCATCTTATGTGCACATGATGCTAGTCCCAACATGGAAAGAGATGCCAGTACCT 963
Db 249 ArgArgLeuThrPheCysHisGlnGlyLeuSerAsnValLeuAspAspProLysSerAla 268
Qy 964 GGAGTCTGCTACCTTTGTTCTCCAGAGAAATTTGATAGATATTCTGGCTATTGGTGGT 1023
Db 269 GlyValAlaThrPheValIleGlnGluGluPheAspArgPheThrGlyTyrTrpTrpCys 288

US-09-976-674-27

Alignment Scores:

Pred. No.: 9,06e-296 Length: 892
Score: 2871.00 Matches: 529
Percent Similarity: 73.7% Conservative: 137
Best Local Similarity: 58.5% Mismatches: 208
Query Match: 51.7% Indels: 30
DB: 2 Gaps: 5

US-10-825-632-2 (1-3120) x US-09-976-674-27 (1-892)

QY	124	TGGAGCGCGCGCAGCATGAACGGCGCGAGCCGCTCCATAGCGCAGCTGGGAGCGTCC	183
DB	17	TrpArgSerPheSerLeuAsnSerGluGlyAlaGluArgMetAlaThrThrGlyThrPro	36
QY	184	GGGCGGGCGCGGGGAGGAATGCAACATGCGCAGCAGCAATGGAACAGACAGCTG	243
DB	37	ThrAlaAspA:gglyAsp-----AlaAlaAla-----	45
QY	244	GGTGTGAGATATTGAAACTGCGGACTGTGAGGAGAAATTTGAATCACAGGATCGGCT	303
DB	46	-----ThrAspAspPro	49
QY	304	AAATTGGAGCCTTTTATGTGAGCGGTATTCTGGAGTCAGCTTAAAGCTGCTGCC	363
DB	50	AlaAlaArg---PheGlnValGlnLysHisSerTrpAspGlyLeuArgSerIleIleHis	68
QY	364	GATACCAAGAAATATCATGTGTACATGATGCTTAAGGCACACATGATTTCTGTTGG	423
DB	69	GlySerArgLysTySerGlyLeuValAsnLysAlaProHisAspPheGlnPheVal	88
QY	424	AAGAGGAATGATCCAGATGAGCCTCATTCACAGCAATCTATTACCTGGCCATGCTGT	483
DB	89	GlnLysThrAspGlnSerGlyProHisSerHisArgLeuTyTrpLeuGlyMetProTy	108
QY	484	GAGACAGAGAAATACACTGTTTATCTGAATTCCTCAAAATCCAAACTATCAATAGCAGCA	543
DB	109	GlySerArgGluAsnSerLeuLeuTySerGluIleProLysLysValArgLysGluAla	128
QY	544	GTCTTAATGCTCTCTTGAAGCCTCTTTTGGATCTTTTTCAGGCAACACTGGACTATGA	603
DB	129	LeuLeuLeuLeuSerTrpLysGlnMetLeuAspHisPheGlnAlaThrProHisHisGly	148
QY	604	ATGTATCTCGAGAGAGAACTATTAAAGAGAAAGAAACCATTCGAAACAGTCGGAATT	663
DB	149	ValTySerArgGluGluGluLeuLeuArgGluArgLysArgLeuGlyValPheGlyIle	168
QY	664	GCTTCTTACGATTATCACCAAGGAAGTGGAACTTTCTGTTTCAAGCCGGTAGTGAATT	723
DB	169	ThrSerTyAspPheHisSerGluSerGlyLeuPheLeuPheGlnAlaSerAsnSerLeu	188
QY	724	TATCAGCTAAAGATGAGGGCCACAGGATTTACGCAACACTTTAAGGCCCAATCTTA	783
DB	189	PheHisCysArgAspGlyGlyLysAsnGlyPheMetValSerProMetLysProLeuGlu	208
QY	784	GTGGAACATGATGTTGCCAACATCAGGATGATCAAAATTTATGCCCGCTGATCCAGAC	843
DB	209	IleLysThrGlnCysSerGlyProArgMetAspProLysIleCysProAlaAspProAla	228
QY	844	TGGATTGCTTTTATACATGCAACGATATTTGGATATCTAACATCGTAACACAGAGAA	903
DB	229	PhePheSerPheIleAsnAsnSerAspLeuTrpValAlaAsnIleGluThrGlyGluGlu	248
QY	904	AGGAGACTCATTATGTGCACAATGAGCTAGCCCAACATGGAAGAGATGCCAGATCAGCT	963
DB	249	ArgArgLeuThrPheCysHisGlnGlyLeuSerAsnValLeuAspAspProLysSerAla	268
QY	964	GGAGTCGCTACCTTTGTTCTCCAAGAGAAATTTGATAGATATTTCTGGCTATTGGTGTG	1023
DB	269	GlyValAlaThrPheValIleGlnGluPheAspArgPheThrGlyTyTrpTrpCys	288
QY	1024	CCAAAAGCTGAAACCACTCCAGTGGTGGT---AAAATTCCTAGATTCTATATGAAGAA	1080

DB	289	ProThrAlaSerTrpGluGlySerGluGlyLeuLysThrLeuArgIleLeuTyGluGlu	308
QY	1081	AATGATGAATCTGAGTGGAAATATTATTCATGTTACATCCCTATGTTGGAAACAGGAGG	1140
DB	309	ValAspGluSerGluValGluValIleHisValProSerProAlaLeuGluGluArgLys	328
QY	1141	GCAGATTTCCTCCGTTATCTTAAACACAGGTACAGCAATCTCAAGTCACCTTTTAAGATG	1200
DB	329	ThrAspSerTyArgTyProArgThrGlySerLysAsnProLysIleAlaLeuLysLeu	348
QY	1201	TCAGAATAATGATTGCTGTAAGGAAGGATCATAGATGTCATAGATAAGAACTAAT	1260
DB	349	AlaGluPheGlnThrAspSerGlnGlyLysIleValSerThrGlnGluLysGluLeuVal	368
QY	1261	CAACCTTTTGGAGATCTTATTGAAGGAGTTGAATATATTGCCAGAGCTGGATGACTCT	1320
DB	369	GlnProPheSerSerLeuPheProLysValGluTyIleAlaArgAlaGlyTrpThrArg	388
QY	1321	GAGGGAATAATGCTTGGTCCATCTCTAGATCGCTCCAGACTCGCTACATAGATG	1380
DB	389	AspGlyLysTyAlaTrpAlaMetPheLeuAspArgProGlnGlnTrpLeuGlnLeuVal	408
QY	1381	TTGATCTCACCTGAATTATTTATCCAGTAGAGATGATGTTATGGAAGGCAGAGACTC	1440
DB	409	LeuLeuProProAlaLeuPheIleProSerThrGlnAsnGluGlnArgLeuAlaSer	428
QY	1441	ATTGAGTCAGTGCCTGATTCGTGACGCCACTTAATTTATCTATGAAGAAACACAGACATC	1500
DB	429	AlaArgAlaValProArgAsnValGlnProTyValValTyGluGluValThrAsnVal	448
QY	1501	TGATATAATATCCATGACATCTTTCATGTTTTCCTCCCAAGTCAC---GAAGAGGAAAT	1557
DB	449	TrpIleAsnValHisAspIlePheTyProPheProGlnSerGluGlyGluAspGluLeu	468
QY	1558	GAGTTATTTTGGCTCTGAATGCAAAACAGAGTTTCCGTATTTATACAAATTCATCT	1617
DB	469	CysPheLeuArgAlaAsnGluCysLysThrGlyPheCysHisLeuTyLysValThrAla	488
QY	1618	ATTTTAAAGAAAGCAATATATAACGATCCAGTGGTGGCTGCTCTCCAAGTGATTC	1677
DB	489	ValLeuLysSerGlnGlyTyAspTrpSerGluProPheSerProGlyGluAspGluPhe	508
QY	1678	AAGTGTCTTAAAGAGGAGATAGCAATTACCATGCTGTAATGCGAGATTTCTTGGCGG	1737
DB	509	LysCysProIleLysGluIleAlaLeuThrSerGlyGluTrpGluValLeuAlaArg	528
QY	1738	CATGGATCTAATATCAAGTTGATGATGATGATGATGATGATGATGATGATGATGATG	1797
DB	529	HisGlySerLysIleTrpValAsnGluGluThrLysLeuValTyPheGlnGlyThrLys	548
QY	1798	GACTCCCTTTAGAGCATCACCTGATGATGATGATGATGATGATGATGATGATGATG	1857
DB	549	AspThrProLeuGluHisIleLeuTyValValSerTyGluAlaAlaGlyGluLeuVal	568
QY	1858	AGGTGACTGACCGTGGCTACTCACATTTCTGCTGCATCAGTCAGTCAGTCAGTCAGT	1917
DB	569	ArgLeuThrThrProGlyPheSerHisSerCysSerMetSerGlnAsnPheAspMetPhe	588
QY	1918	ATAAGTAAGTATAGTAACCAAGAAATCCACTGTGTGCTCCCTTTTCAAGCTATCAAGT	1977
DB	589	ValSerHisTySerSerValSerThrProCysValHisValTyLysLeuSerGly	608
QY	1978	CTGAAGATGACCAACTTGCACAAAGAAATTTTGGGCCACCATTTTGGATTCAGCA	2037
DB	609	ProAspAspAspProLeuHisLysGlnProArgPheTrpAlaSerMetMetGluAla	628
QY	2038	GGTCTCTTCTGCTATCTCTCCAGAAATTTCTCTTTTGAAGTACTACTGATTT	2097
DB	629	SerCysProProAspTyValProProGluIlePheHisPheHisThrArgSerAspVal	648
QY	2098	ACATTGTATGGGATGCTTACAGGCTCATGATCTACAGCTCGGAAAGAAATATCTACT	2157

Db	649	ArgLeuTyrGlyMetIleTyrLysProHisAlaLeuGlnProGlyLysLysHisProThr	668
Qy	2158	GTGCTGTCATATATGGTCTCTCAGGTGCAGTTGGTGAATAATCGGTTTAAAGGAGTC	2217
Db	669	ValLeuPheValTyrGlyGlyProGlnValGlnLeuValAsnAsnSerPheLysGlyIle	698
Qy	2218	AAGTATTTCGCTTGAATACCCCTAGCCTCTCTAGGTATGTGGTTGTAGTCATAGACAAC	2277
Db	689	LysTyrLeuArgLeuAsnThrLeuAlaSerLeuGlyTyrAlaValValIleAspGly	708
Qy	2278	AGGGATCCTGTACCGAGGCGCTTAAATTTGAAGCGCCTTTAAATATATAAAATGGGTCAA	2337
Db	709	ArgGlySerCysGlnArgGlyLeuArgPheGluGlyAlaLeuLysAsnGlnMetGlyGln	728
Qy	2338	ATAGAAATTCACCATCAGGTGGAAGCACTCCATATCTAGCTTCTCGATATGATTTCAAT	2397
Db	729	ValGluLeuGlnAspGlnValGluGlyLeuGlnPheValAlaGluLysTyrGlyPheIle	748
Qy	2398	GACTAGATCGTGTGGGCATCCAGCGCTGCTCTATGGAGGATACCTCTCCCTGATGCCA	2457
Db	749	AspLeuSerArgValAlaIleHisGlyTyrSerTyrGlyGlyPheLeuSerLeuMetGly	768
Qy	2458	TTAATGCAGAGTCAGATATCTTCAGGGTTGCTATTGCTGGGGCCCGCCAGTCACCTGTGG	2517
Db	769	LeuIleHisLysProGlnValPheLysValAlaIleAlaGlyAlaProValThrValTyrP	788
Qy	2518	ATCTTCATGATACAGATACACGGAACGTTATATGGTTCACCTGACACGAAATGACAG	2577
Db	789	MetAlaTyrAspThrGlyTyrThrGluArgTyrMetAspValProGluAsnAsnGlnHis	808
Qy	2578	GGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAATTCCTCTGACCAAAATCGT	2637
Db	809	GlyTyrGluAlaGlySerValAlaLeuHisValGluLysLeuProAsnGlnProAsnArg	828
Qy	2638	TTACTGCTCTTACATGGTTTCCTGGATGAGATGTCCATTTTGCAATACCAATATATTA	2697
Db	829	LeuLeuIleLeuHisGlyPheLeuAspGluAsnValHisPhePheHisThrAsnPheLeu	848
Qy	2698	CTGAGTCTTTTGTAGTGAGGCTGGAAAGCCATATGATTTACAGATCTATCTCCTCAGGAGA	2757
Db	849	ValSerGlnLeuIleAlaGlyLysProTyrGlnLeuGlnIleTyrProAsnGluArg	868
Qy	2758	CACAGCATACAGTTCTCGAATCGGAGAACATTATGAATCTCTTTTGACATACCTT	2817
Db	869	HisSerIleAlaGlyCysProGluSerGlyGluHisTyrGluValThrLeuLeuHisPheLeu	888
Qy	2818	CAAGAAACCTT 2829	
Db	889	GlnGluTyrLeu 892	
RESULT 9			
US-09-976-674-3			
; Sequence 3, Application US/09976674			
; Patent No. 6844180			
; GENERAL INFORMATION:			
; APPLICANT: Qi, Steve			
; APPLICANT: Akinsanya, Karen			
; APPLICANT: Riviere, Pierre			
; APPLICANT: Junien, Jean-Louis			
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV			
; FILE REFERENCE: 70669			
; CURRENT APPLICATION NUMBER: US/09/976,674			
; CURRENT FILING DATE: 2001-10-12			
; PRIOR APPLICATION NUMBER: US 60/240,117			
; PRIOR FILING DATE: 2000-10-12			
; NUMBER OF SEQ ID NOS: 61			
; SOFTWARE: PatentIn version 3.1			
; SEQ ID NO 3			
; LENGTH: 863			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
US-09-976-674-3			

Alignment Scores:			
Pred. No.:	1,13e-295	Length:	863
Score:	2870.00	Matches:	517
Percent Similarity:	77.5%	Conservative:	134
Best Local Similarity:	61.5%	Mismatches:	187
Query Match:	51.7%	Indels:	2
DB:	2	Gaps:	2
US-10-825-632-2 (1-3120) x US-09-976-674-3 (1-863)			
Qy	316	TTTTATGTTGACCGGTATCTCTGAGTCAGCTTAAAAAGCTGCTTCCGATACCAGAAA	375
Db	24	PheGlnValGlnLysHisSerTrpAspGlyLeuArgSerIleIleHisGlySerArgLys	43
Qy	376	TATCATGGCTCATGATGGCTAAGGCCACCATGATTTTCATGTTGTGAAGAGGAATGAT	435
Db	44	TyrSerGlyLeuIleValAsnLysAlaProHisAspPheGlnPheValGlnLysThrAsp	63
Qy	436	CCAGATGCACCTCATTCAGACAGAAATCTATTACCTTCCCATGTCCTCGTGAGAACAGAA	495
Db	64	GluSerGlyProHisSerHisArgLeuTyrTyrLeuGlyMetProTyrGlySerArgGlu	83
Qy	496	AATACACTGTTTATTTCTGAAATTCCTCAAACTATCAATAGACGACAGCTCTTAATGCTC	555
Db	84	AsnSerLeuLeuTyrSerGluIleProLysLysValArgLysGluAlaLeuLeuLeu	103
Qy	556	TCCTTGAAGCCTCTTTTGGATCTTTTTCAGCAACACTGACACTATGGAATGTATTCGCA	615
Db	104	SerTrpLysGlnMetLeuAspPheGlnAlaThrProHisGlyValTyrSerArg	123
Qy	616	GAAGAAGAACTATTAAAGAGAAAGAACGCAATGGAACAGTCGGAATTTCTTACGAT	675
Db	124	GluGluGluLeuLeuLeuArgGluArgLysLeuGlyValPheGlyIleThrSerTyrAsp	143
Qy	676	TATCAACCAAGNAGTGGNACATTTCTTTCACGCGGTAGTGGATTTATATACGTAATAA	735
Db	144	PheHisSerGluSerGlyLeuPheLeuPheGlnAlaSerAsnSerLeuPheHisCysArg	163
Qy	736	GATGAGGCGCCAAAGGATTTACGCAACAACCTTTAAGGCCCAATCTAGTGGAAACTAGT	795
Db	164	AspGlyGlyLysAsnGlyPheMetValSerProMetLysProLeuGluIleLysThrGln	183
Qy	796	TGTCCCAACATACGATCGATCCAAAATTTATGCCCGCTGATCCAGACTGGATGCTTTT	855
Db	184	CysSerGlyProArgMetAspProLysIleCysProAlaAspProAlaPheSerPhe	203
Qy	856	ATACATAGCAACGATATTGGATATCTAACTCTAACTCTAACTCTAACTCTAACTCT	915
Db	204	IleAsnAsnSerAspLeuTrpValAlaAsnIleGluThrGlyGluGluArgLeuThr	223
Qy	916	TATGTGCAATGAGCTAGCCCAACATGGAAGAAGATCCAGATCAGCTGAGTGCCTACC	975
Db	224	PheCysHisGlnGlyLeuSerAsnValLeuAspAspProLysSerAlaGlyValAlaThr	243
Qy	976	TTTGTCTCCAAAGAAAGATTTGATAGATATTTCTGGCTATTGGTGGTGTCCAAAGCTGAA	1035
Db	244	PheValIleGlnGluGluPheAspArgPheThrGlyTyrTrpTrpCysProThrAlaSer	263
Qy	1036	ACAACTCCAGTGGTGGT---AAAATCTTAAAGATTTCTATATGAAGAAAATGATGATCT	1092
Db	264	TrpGluGlySerGluGlyLeuLysThrLeuArgIleLeuTyrGluGluValAspGluSer	283
Qy	1093	GAGTGGAAATTTATCATGTTTACATCCCTATGTTGGAAACAAAGAGGCGGATTCATTC	1152
Db	284	GluValGluValIleHisValProSerProAlaLeuGluArgLysThrAspSerTyr	303
Qy	1153	CGTTATCTTAAACAGGTACACAAATCTCAAGTCACCTTTTAAAGATGTTCAGAAATAATG	1212
Db	304	ArgTyrProArgThrGlySerLysAsnProLysIleAlaLeuLysLeuAlaGluPheGln	323
Qy	1213	ATTGATGCTGAAGAAAGGATCATAGATGTCTAGATAGGAACCTTAATTCACCTTTGAG	1272
Db	324	ThrAspSerGlnGlyLysIleValSerThrGlnGluLysGluLeuValGlnProPheSer	343


```
QY 1273 ATTCTATTTGAAGGAGTTGAATATATTGCGAGAGCTGGATGGACTCTCGAGGCAAAATAT 1332
DB |||||
DB 344 SerLeuPheProLysValGluTyrIleAlaArgAlaGlyTrpThrArgAspGlyLysTyr 363
QY 1333 GCTTGGTCCATCTACTAGATCGCTCCAGACTCGCTACAGACTAGTGTGTATCTCACCT 1392
DB |||||
DB 364 AlaTrpAlaMetPheLeuAspArgProGlnGlnTrpLeuGlnValLeuLeuProPro 383
QY 1393 GAATTTATTTATCCAGTAGAAGATGATGTTATGGAAGAGGAGAGACTCATTTGAGTCAGTG 1452
DB |||||
DB 384 AlaLeuPheIleProSerThrGluAsnGluGluArgLeuAlaSerAlaArgAlaVal 403
QY 1453 CCTGATCTCTGAGCCCACTAATTTATCTATGAAGAACAACAGACATCTGGTAATATC 1512
DB |||||
DB 404 ProArgAsnValGlnProTyrValValTyrGluGluValThrAsnValTrpIleAsnVal 423
QY 1513 CATGACATCTTTTCATGTTTTCGCCCAAGTCAC---GAAGAGGAATTCAGTTTATTTT 1569
DB |||||
DB 424 HisAspIlePheTyrProPheProGlnSerGluGlyGluAspGluLeuCysPheLeuArg 443
QY 1570 GCCTCTGAATGCAAAACAGGTTTCCGTCATTATACAAATTAACATCTATTTTAAAGGAA 1629
DB |||||
DB 444 AlaAsnGluCysLysThrGlyPheCysHisLeuTyrLysValThrAlaValLeuLysSer 463
QY 1630 AGCAATATAAACCATCCAGTGTGGGCTCGCTCCAGTGATTTCAAGTGTCTATC 1689
DB |||||
DB 464 GlnGlyTyrAspTrpSerGluProPheSerProGlyGluAspGluPheLysCysProIle 483
QY 1690 AAAGAGGAGATGACAAATPACAGTGTGTAATGGGAAGTCTTGGCCGGATGATCTAAT 1749
DB |||||
DB 484 LysGluGluIleAlaLeuThrSerGlyGluTrpGluValLeuAlaArgHisGlySerLys 503
QY 1750 ATCCAAGTTGATGAAGTCAGAGGCTGGTATATTTGAAGGCAACCAAGACTCCCTTTA 1809
DB |||||
DB 504 IleTrpValAsnGluGluThrLysLeuValTyrPheGlnGlyThrLysAspThrProLeu 523
QY 1810 GAGCATCACTGTAGTACGTACGTAAATCTGGAGGTGACAGGTCACAGGTCAGTCAGC 1869
DB |||||
DB 524 GluHisHisLeuTyrValValSerTyrGluAlaAlaGlyGluIleValArgLeuThrThr 543
QY 1870 CTGGCTACTCACATCTCTGTCATCAGTCAGTCAGTCTCTTCTTATTAAGTAAGTAT 1929
DB |||||
DB 544 ProGlyPheSerHisSerCysSerMetSerGlnAsnPheAspMetPheValSerHisTyr 563
QY 1930 AGTAACCAAGAAATCCACACTGTGTCTCTTACCAAGCTATCAAGTCTCTGAAGATGAC 1989
DB |||||
DB 564 SerSerValSerThrProProCysValHisValTyrLysLeuSerGlyProAspAspAsp 583
QY 1990 CCAACTTGCAAAACAAAGGAATTTTGGGCCACCAATTTTGGATTCAGCAGGTCCTCTTCT 2049
DB |||||
DB 584 ProLeuHisLysGlnProArgPheTrpAlaSerMetMetGluAlaAlaSerCysProPro 603
QY 2050 GACTATATCTCTCCAGAAATTTCTCTTTTGAAGTACTACTCGATTTACATTTGATGGG 2109
DB |||||
DB 604 AspTyrValProProGluIlePheHisPheHisThrArgSerAspValArgLeuTyrGly 623
QY 2110 ATGCTCTACAGCCTCATGATCTACAGCTGGGAAGAAATATCTTACTGTGCTGTTCATA 2169
DB |||||
DB 624 MetIleTyrLysProHisAlaLeuGlnProGlyLysLysHisProThrValLeuPheVal 643
QY 2170 TATGCTGTGCTCAGTGCAGTCTGGTGAATATCGGTTTAAAGGAGTCAGATTTTCCGC 2229
DB |||||
DB 644 TyrGlyGlyProGlnValGlnLeuValAsnAsnSerPheLysGlyIleLysTyrLeuArg 663
QY 2230 TTGAATACCTAGCTCTCTAGGTTATGTGTTGTAGTATAGACAACAGGGGATCTCTGT 2289
DB |||||
DB 664 LeuAsnThrLeuAlaSerLeuGlyTyrAlaValValIleAspGlyArgGlySerCys 683
QY 2290 CACCGAGGCTTTAAATTTGAAGGCGCTTTTAAATATATAATGGGTCAAAATAGAAATGAC 2349
DB |||||
DB 684 GlnArgGlyLeuArgPheGluGlyAlaLeuLysAsnGlnMetGlyGlnValGluIleGlu 703
```

```
QY 2350 GATCAGTGGAAAGGACTCCAATATATCTAGCTTTCTCGATATGATTTTCATTGATCTAGATCGT 2409
DB |||||
DB 704 AspGlnValGluGlyLeuGlnPheValAlaGluLysTyrGlyPheIleAspLeuSerArg 723
QY 2410 GTGGGATCCACGGCTGTCTTATGGAGGATACCTCTCCCTGATGGCATTAATTCAGAGG 2469
DB |||||
DB 724 ValAlaIleHisGlyTrpSerTyrGlyPheLeuSerLeuMetGlyLeuIleHisLys 743
QY 2470 TCAGATATCTTCAGGGTGTCTATTCCTGGGGCCCGAGTCACCTCTGTGATCTTCTATGAT 2529
DB |||||
DB 744 ProGlnValPheLysValAlaIleAlaGlyAlaProValThrValTrpMetAlaTyrAsp 763
QY 2530 ACAGGATACACGAAGCTTATATGGGTCACTCCCTGACCAAGATGAACAGGGCTATTACTTA 2589
DB |||||
DB 764 ThrGlyTyrThrGluArgTyrMetAspValProGluAsnAsnGlnHisGlyTyrGluAla 783
QY 2590 GGATCTGTGGCCATCAGACAGAAAGTTCCCTCTGAAACCAATCGTTTACTGCTCTTA 2649
DB |||||
DB 784 GlySerValAlaLeuHisValGluLysLeuProAsnGluProAsnArgLeuLeuIleLeu 803
QY 2650 CATGGTTTCTCGATGAGAAATGTCATTTTGGACATACCATGATATTACTGAGTTTTAA 2709
DB |||||
DB 804 HisGlyPheLeuAspGluAsnValHisPhePheHisThrAsnPheLeuValSerGlnLeu 823
QY 2710 GTGAGGCTGGAAAGCCATATGATTTACAGATCTATCTCTCAGGAGACACAGCATAGA 2769
DB |||||
DB 824 IleArgAlaGlyLysProTyrGlnLeuGlnIleTyrProAsnGluArgHisSerIleArg 843
QY 2770 GTTCTGTAATCGGAGAACATTAATCACTGTCATCTTTTGGCACTACCTTCAAGAAACCTT 2829
DB |||||
DB 844 CysProGluSerGlyGluHisTyrGluValThrLeuLeuHisPheLeuGlnGluTyrLeu 863

RESULT 10
US-09-976-674-33
; Sequence 33, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCES: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 33
; LENGTH: 879
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-33

Alignment Scores:
Pred. No.: 1,66e-290 Length: 879
Score: 2821.50 Matches: 522
Percent Similarity: 72.7% Conservative: 135
Best Local Similarity: 57.7% Mismatches: 204
Query Match: 50.8% Indels: 43
DB: 2 Gaps: 6

US-10-825-632-2 (1-3120) x US-09-976-674-33 (1-879)
QY 124 TGGAGGCGGCGCAGCATGTAAGCGCGCAGCGCCGCTCCATAGCGCACGTCGGGACGTC 183
DB |||||
DB 17 TrpArgSerPheSerLeuAsnSerGluGlyAlaGluArgMetAlaThrThrGlyThrPro 36
QY 184 GCGCGGCGCGCGGGAAGAAATGCAACATGGCAGCAGCAATGGAAACAGACAGCTG 243
DB |||||
DB 37 ThrAlaAspArgGlyAsp-----AlaAlaAla----- 45
```

Qy	244	GGTGTGAGATATTTGAAACTCGGACTGTGAGGAGATATTTGAATCACAGGATCGGCCT	303	Db	389	AspGlyLysTyrAlaTrpAlaMetPheLeuAspArgProGlnGlnTrpLeuGlnLeuVal	408
Db	46	-----ThrAspAspPro	49	Qy	1381	TTGATCTCACTGAATTTATTTATCCAGTAGAAGATGATGTTATGGAAGGAGAGACTC	1440
Qy	304	AAATTGGAGCCTTTTATGTTGAGCGGTATTCCTGGAGTCAGCTTAAAGAGCTCTTGGCC	363	Db	409	LeuLeuProProAlaLeuPheIleProSerThrGluAsnGluGlnArgLeuAlaSer	428
Db	50	AlaAlaArg---PheGlnValGlnLysHisSerTrpAspGlyLeuArgSerIleHis	68	Qy	1441	ATTGAGTCAGTCCCTGATCTGTGACGCCACTAATTTATCTATGAAAGAACACAGACATC	1500
Qy	364	GATACCAAGAAATATCATCGCTACATGATGGCTAAGGACACACATGATTTTCATGTTGTG	423	Db	429	AlaArgAlaValProArgAsnValGlnProTyrValValTyrGluGluValThrAsnVal	448
Db	69	GlySerArgLysTyrSerGlyLeuIleValAsnLysAlaProHisAspPheGlnPheVal	88	Qy	1501	TGGATAAATATCCATGACATCTTTTCATGTTTTTCCCAAGATCAC---GAAGAGGAAAT	1557
Qy	424	AAGGAATATCCAGATCGACCTCATTCAGACAGACTATTAATCTTCCCTGACATCTCGGT	483	Db	449	TrpIleAsnValHisAspIlePheTyrProPheProGlnSerGluGlyGluAspGluLeu	468
Db	89	GlnLysThrAspGluSerGlyProHisSerHisArgLeuTyrTyrLeuGlyMetProTyr	108	Qy	1558	GAGTTTATTTTGGCTCTGAAATGCAAAACAGGTTTCCGTCATTTATACAAAATTCATCT	1617
Qy	484	GAGAACAGAGAAATACATCTGTTTATCTGAAATTTCCAAACTATCAATAGACAGCA	543	Db	469	CysPheLeuArgAlaAsnGluCysLysThrGlyPheCysHisLeuTyrLysValThrAla	488
Db	109	GlySerArgGluAsnSerLeuLeuTyrSerGluIleProLysLysValArgLysGluAla	128	Qy	1618	ATTTTAAAGGAAAGCAAAATATAAACGATCCAGTGGTGGCTGCTCCCAAGTGATTC	1677
Qy	544	GTCTTAATGCTCTCTGGAAGCCTCTTTTGGATCTTTTCAGGCAACCTCGACTATGGA	603	Db	489	ValLeuLysSerGlnGlyTyrAspTrpSerGluProPheSerProGlyGluAspGluPhe	508
Db	129	LeuLeuLeuLeuSerTrpLysGlnMetLeuAspHisPheGlnAlaThrProHisGly	148	Qy	1678	RAGTCTCTATCAAGAGGATAGCAATACAGTGGTGAATGGGAAGTTCTTGGCCGG	1737
Qy	604	ATGATTTCTCAGAGAAGAACTATTAAAGAGAAGAAAACGCAATGGAACTCGGAAT	663	Db	509	LysCysProIleLysGluGluIleAlaLeuThrSerGlyGluTrpGluValLeuAlaArg	528
Db	149	ValTyrSerArgGluGluGluLeuLeuLeuArgGluArgLysArgLeuGlyValPheGlyIle	168	Qy	1738	CATGATCTAATATCAAGTTGATGAAGTCAGAAGCTGGTATATTTTGAAGGCACCAAA	1797
Qy	664	GCTTCTTACGATATACCAAGGAAGTGGAAATTTCTGTTTCAAGCCGGTAGTGAAAT	723	Db	529	HisGlySer-----LysGlyThrLys	535
Db	169	ThrSerTyrAspPheHisSerGluSerGlyLeuPhePheGlnAlaSerAsnSerLeu	188	Qy	1798	GACTCCCTTTAGACATCACTGTAGTAGTCAGTTACGTAAATCCTGGAGAGTGACA	1857
Qy	724	TATCAGCTAAAGATGAGGGCCACAAAGGATTTTACCAACACCTTTAAGGCCCAATCTA	783	Db	536	AspTrpProLeuGluHisLeuTyrValSerTyrGluAlaGlyGluIleVal	555
Db	189	PheHisCysArgAspGlyGlyLysAsnGlyPheMetValSerProMetLysProLeuGlu	208	Qy	1858	AGCGTACTGACCGTGGTACTCACATCTTGTGTCAGTCAGTCAGCTGTGACTCTTT	1917
Qy	784	GTGGAATCTGTTCCCAACATACGATGGATCCAAATATATGCCCGCTGATCCAGAC	843	Db	556	ArgLeuThrThrProGlyPheSerHisSerCysSerMetSerGlnAsnPheAspMetPhe	575
Db	209	IleLysThrGlnCysSerGlyProArgMetAspProLysIleCysProAlaAspProAla	228	Qy	1918	ATAAGTAAGTATAGTAACCAAGAAATCCACACTGTGTGTCCTTTTCAAGCTATCAAGT	1977
Qy	844	TGGATGCTTTTATACATACCAAGATTTTGGATATCTAATCATCTTAACCAAGAGAAGA	903	Db	576	ValSerHisTyrSerSerValSerThrProCysValHisValTyrLysLeuSerGly	595
Db	229	PhePheSerPheIleAsnAsnSerAspLeuTrpValAlaAsnIleGluThrGlyGluGlu	248	Qy	1978	CCTGAAGATGACCCCAACTTGCAAAACAAAGAAATTTTGGGCCACCATTTTGGATTCAGCA	2037
Qy	904	AGGAGCTACATTTATGTCACATGACCTAGCCACATGGAAGAGATCCAGATCAGCT	963	Db	596	ProAspAspAspProLeuHisLysGlnProArgPheTrpAlaSerMetGluAlaAla	615
Db	249	ArgArgLeuThrPheCysHisGlnGlyLeuSerAsnValLeuAspPheProLysSerAla	268	Qy	2038	GGTCTCTTCTCGACTACTCTCCAGAAATTTTCTCTTTTGAAGTACTACTGGATTT	2097
Qy	964	GGAGTCGCTACTTTTGTCTCCAAAGAAATTTGATAGATATTCGCTATTTCGTGGT	1023	Db	616	SerCysProProAspTyrValProProGluIlePheHisPheHisThrArgSerAspVal	635
Db	269	GlyValAlaThrPheValIleGlnGluPheAspArgPheThrGlyTyrTrpPheCys	288	Qy	2098	ACATTTGATGGGATGCTCTACAAAGCTCATGATCTACAGCTCGAAAGAAATATCCTACT	2157
Qy	1024	CCAAAAGCTGAACAACTCCAGTGGTGGT---AAAATCTTGAATCTTATATGAAGA	1080	Db	636	ArgLeuTyrGlyMetIleTyrLysProHisAlaLeuGlnProGlyLysLysHisProThr	655
Db	289	ProThrAlaSerTrpGluGlySerGluGlyLeuLysThrLeuArgIleLeuTyrGluGlu	308	Qy	2158	GTGCTGTCATATATGGTGGTCTCAGCTGAGTTGGTGAATAATCGTTTAAAGGAGTC	2217
Qy	1081	AATGATGAATCTGAGGTGGAATTTATTCATGTTACATCCCTATGTTGGAACAAAGAGG	1140	Db	656	ValLeuPheValTyrGlyProGlnValGlnLeuValAsnSerPheLysGlyIle	675
Db	309	ValAspGluSerGluValGluValIleHisValProSerProAlaLeuGluArgLys	328	Qy	2218	AAGTATTTCCGTTCTGATACCTAGCTCTCTAGTATATGGTTGTAGTAGTAGACAAAC	2277
Qy	1141	GCAGATTCATTCGTTATCTTAAACACAGGTACAGCAATCTTAAAGTCACTTTAAGATG	1200	Db	676	LysTyrLeuArgLeuAsnThrLeuAlaSerLeuGlyTyrAlaValValIleAspGly	695
Db	329	ThrAspSerTyrArgTyrProArgThrGlySerLysAsnProLysIleAlaLeuLysLeu	348	Qy	2278	AGGGGATTCCTGTCAACCGAGGCTTAAATTTGAGCGCTTTAAATATAAATGGGCAA	2337
Qy	1201	TCAGAAATAATGATTGCTGGAAGGAGATCATAGATGCTCATAGATAGGAACATAAT	1260	Db	696	ArgGlySerCysGlnArgGlyLeuArgPheGluGlyAlaLeuLysAsnGlnMetGlyGln	715
Db	349	AlaGluPheGlnThrAspSerGlnGlyLysIleValSerThrGlnGluLysGluLeuVal	368	Qy	2338	ATAGAAATGACATCAGGTGGAAGGATCCCAATATCTAGCTTCTCGATATGATTCAT	2397
Qy	1261	CAACCTTTTGAGTTCTTATGGAAGGAGTTGAATATATTCGAGAGCTGATGAGCTCT	1320	Db	716	ValGluIleGluAspGlnValGluGlyLeuGlnPheValAlaGluLysTyrGlyPheIle	735
Db	369	GlnProPheSerSerLeuPheProLysValGluTyrIleAlaArgAlaGlyTrpThrArg	388	Qy	2398	GACTTAGATGCTGTGGGCATCCACGGCTGCTCTATGAGGAGATCCTCTCCCTGATGCA	2457
Qy	1321	GAGGGAANAATATGCTTGGTCCATCTACTAGATCGCTCCCGAGATGCTACAGATAGTG	1380				

Db 736 AspLeuSerArgValAlaIleHisGlyTrpSerTyrglyGlyPheLeuSerLeuMetGly 755
Qy 2458 TTAATGAGAGGTGAGATATCTTACAGGTTGCTATTGCTGGGCCCCAGTCACTCTGNG 2517
Db 756 LeuIleHisLysProGlnValPheLeValaIleAlaGlyAlaProValThrValTrp 775
Qy 2518 ATCTTCATATACAGGATACAGGACGTTATATATGGGTCACTGACCAAGTAAACAG 2577
Db 776 MetAlaTyrrpThrGlyTyrrThrGluArgTyrrMetAspValProGluAsnAsnGlnHis 795
Qy 2578 GGCTATTACTTAGGATCTGTGGCCATGCAAGCAGAAAAGTTCCTCTGCAACCAATCGT 2637
Db 796 GlyTyrrGluAlaGlySerValAlaLeuHisValGluLysLeuProAsnGluProAsnArg 815
Qy 2638 TTACTGCTTTACATGTTTCTGGATGAGAAATGTCATTTGGCATACATCCAGTATATTA 2697
Db 816 LeuLeuIleLeuHisGlyPheLeuAspGluAsnValHisPhePheHisThrAsnPheLeu 835
Qy 2698 CTGAGTTTTTAGTGAGGCTGGNAAGCCATATGATTACAGATCTATCTCTCAGGAGAGA 2757
Db 836 ValSerGlnLeuIleArgAlaGlyLysProTyrrGlnLeuGlnIleTyrrProAsnGluArg 855
Qy 2758 CACAGCATTAAGAGTTCTGGAATCGGAGAACATTATGAATGCACTCTTTTGCACTACCTT 2817
Db 856 HisSerIleArgCysProGluSerGlyGluHisTyrrGluValThrLeuLeuHisPheLeu 875
Qy 2818 CAAGAAAACCTT 2829
Db 876 GlnGluTyrrLeu 879

RESULT 11
US-09-976-674-35
; Sequence 35, Application US/09976674
; Patent No. 684180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 35
; LENGTH: 879
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-35

Alignment Scores:
Pred. No.: 1.66e-290 Length: 879
Score: 2821.50 Matches: 522
Percent Similarity: 72.7% Conservative: 135
Best Local Similarity: 57.7% Mismatches: 204
Query Match: 50.8% Indels: 43
DB: 2 Gaps: 6

US-10-825-632-2 (1-3120) x US-09-976-674-35 (1-879)

Qy 124 TGGAGCGCGGAGCATGTAACGGCGGAGCCGCTCCATAGCGCAGCTCGGACGGTCC 183
Db 17 TrpArgSerPheSerLeuAsnSerGluGlyAlaGluArgMetAlaThrThrGlyThrPro 36
Qy 184 GGGCGGCGCGGGGGAAGAAATGCAATCGCAGCAGCAATGGAACAGACAGCTG 243
Db 37 ThrAlaAspArgGlyAsp-----AlaAlaAla----- 45
Qy 244 GGTGTTGAGATATTGAAACTGCGGACTGTGAGGAGAGATATTGAATCACAGGATCGGCT 303

Db 46 -----ThraspAspPro 49
Qy 304 AAATTGGAGCCTTTTATGTTGAGCGTATTCCTGGAGTCAGCTTAAAAAGCTCTTGCC 363
Db 50 AlaAlaArg---PheGlnValGlnLysHisSerTrpAspGlyLeuArgSerIleIleHis 68
Qy 364 GATACACAGAAAATATCATGGCTACATGATGCTAAGCCACACATGATTTCATGTTGTG 423
Db 69 GlySerArgLysTyrrSerGlyLeuIleValAsnLysAlaProHisAspPheGlnPheVal 88
Qy 424 AAGAGGAATGATCCAGATGGAGCTCATTCAGACAGATCTATTACCTTGCCATGCTCTGGT 483
Db 89 GlnLysThrAspGluSerGlyProHisSerHisArgLeuTyrrTyrrLeuGlyMetProTyrr 108
Qy 484 GAGAACAGAAAATACACTGTTTATTCCTGAATTCCTCAAACTATCAATAGACGACGA 543
Db 109 GlySerArgGluAsnSerLeuLeuTyrrSerGluIleProLysLysValArgLysGluAla 128
Qy 544 GTCTTAATCTCTCTTGGAGGCTCTTTTGGATCTTTTTCAGGCAACACTGAGCTATGGA 603
Db 129 LeuLeuLeuLeuSerTrpLysGlnMetLeuAspHisPheGlnAlaThrProHisGly 148
Qy 604 ATGTATTCGAGAGAAAGAACTATTAAAGAGAAAGAAACGATTTGGAACAGTCGGAATT 663
Db 149 ValTyrrSerArgGluGluGluLeuLeuArgGluArgLysArgLeuGlyValPheGlyIle 168
Qy 664 GCTTCTTACGATTTATCACCAAGGAAGTGGAAACATTCTGTTTCAAGCGGTGATGGAATT 723
Db 169 ThrSerTyrrAspPheHisSerGluSerGlyLeuPheLeuPheGlnAlaSerAsnSerLeu 188
Qy 724 TATCAGCTAAAAGATGGAGGCCACAGGATTTACGCAACCACTTTAAAGCCCCAATCTA 783
Db 189 PheHisCysArgAspGlyGlyLysAsnGlyPheMetValSerProMetLeuLeuGlu 208
Qy 784 GTGGAACCTAGTTGTCCCAACATACGATGGATGATCCAAATTTATGCCCGCTGATCCAGAC 843
Db 209 IleLysThrGlnCysSerGlyProArgMetAspProLysIleCysProAlaAspProAla 228
Qy 844 TGGATTGCTTTTATATACATAGCAACGATTTTGGATATCTAAACATCGTAAACAGAGAA 903
Db 229 PhePheSerPheIleAsnAsnSerAspLeuTrpValAlaAsnIleGluThrGlyGluGlu 248
Qy 904 AGGAGACTCACTTATGTGCACATAGCTAGCTAGCCACATCGAAGAGATGCCAGATCAGCT 963
Db 249 ArgArgLeuThrPheCysHisGlnGlyLeuSerAsnValLeuAspAspProLysSerAla 268
Qy 964 GGAGTCGCTACCTTTGTTCTCCAAGAAGAATTGATAGATATCTGGCTATTTGGTGTGT 1023
Db 269 GlyValAlaThrPheValIleGlnGluPheAspArgPheThrGlyTyrrTrpTrpCys 288
Qy 1024 CCAAAAAGCTGAACAACTCCAGTGGTGT---AAAATCTTAGAATCTTATATGAAGA 1080
Db 289 ProThrAlaSerTrpGluGlySerGluGlyLeuLysThrLeuArgIleLeuTyrrGluGlu 308
Qy 1081 AATGATGAATCTGAGGTGGAATTTATCATGTTATCATCCCTATGTTGGAACAGAGAGG 1140
Db 309 ValAspGluSerGluValGluValIleHisValProSerProAlaLeuGluGluArgLys 328
Qy 1141 GCAGATTTCATTCGTTATCTTAAACAGGTACAGCAAAATCTTAAAGTCACTTTTAAAGAT 1200
Db 329 ThrAspSerTyrrArgTyrrProArgThrGlySerLysAsnProLysIleAlaLeuLysLeu 348
Qy 1201 TCAGAAATATGATGATGCTGAAGAGGATCATAGATGTCTATAGATGTCATAGATAAGAACTAAT 1260
Db 349 AlaGluPheGlnThrAspSerGlnGlyLysIleValSerThrGlnGluLysGluLeuVal 368
Qy 1261 CAACCTTTGAGATCTTATTGAGAGGTGGAATGATATATTGGCAGAGCTGGATGGACTCCT 1320
Db 369 GlnProPheSerSerLeuPheProLysValGluTyrrIleAlaArgAlaGlyTrpThrArg 388
Qy 1321 GAGGGAATAATATGCTTTGGTCCATCTCTATAGATCGTCCAGACTCGGCTACAGATAGTG 1380
Db 389 AspGlyLysTyrrAlaTrpAlaMetPheLeuAspArgProGlnGlnTrpLeuGlnLeuVal 408

Qy	1381	TTGATCTCACTGAAATTTATCCAGTAGAAGATGTTATGGAAGCCAGACTC	1440
Dd	409	LeuLeuProAlaLeuPheLeuProSerThrGluAenGluGluGluGluAeUaSer	428
Qy	1441	ATTGAGTCAGTGGCTGATCTGTGACGCCACTAATTTATGAAAGAACACAGATC	1500
Dd	429	AlaArgAlaValProArgAenValGlnProTyrValValTyrGluGluValThrAenVal	448
Qy	1501	TGGATAATATCCATGACATCTTTTCATGTTTTTCCCAAGTCAC---GAAGAGGAAT	1557
Dd	449	TrpIleAenValHisAenPheTyrProPheProGlnSerGluGluGluAenGluLeu	468
Qy	1558	GAGTTTATTTTGGCTCTGAATGCAAAACAGGTTTCGGTCATTTATACAAATATACATCT	1617
Dd	469	CysPheLeuAenGluAenGluCysLeuThrGlyPheCysHisLeuTyrLysValThrAla	488
Qy	1618	ATTTTAAAGGAAGCAATATAACGATCCAGTGGTGGCTCCCTGCTCCCAAGTATTC	1677
Dd	489	ValLeuLysSerGlnGlyTyrAspTyrPheSerGluProPheSerProGlyGluAenGluPhe	508
Qy	1678	AGTGTCTCTATCAAGGAGGATAGCAATTACCAGTGGTGAATGGAGTTCTTGGCGG	1737
Dd	509	LysCysProLeuLysGluGluLeuLeuLeuThrSerGlyLeuTrpGluValLeuAlaArg	528
Qy	1738	CATGGATCTAATATCCAAAGTTGATGAAGTCAGAGCTCGTATATTTTGAAGGCCCAA	1797
Dd	529	HisGlySer-----LysGlyThrLys	535
Qy	1798	GACTCCCTTTAGACATCACTGTAGTAGTACGTATTAATCTCGGAGAGGTGACA	1857
Dd	536	AspThrProLeuGluHisLeuTyrValValSerTyrGluAlaLeuGlyGluLeuVal	555
Qy	1858	AGGTGATGACCGTGGCTACTCACATTTCTGTGTCATGTCAGTCAGCAGCTGTGACTCTTT	1917
Dd	556	ArgLeuThrThrProGlyPheSerHisSerCysSerMetSerGlnAenPheAspMetPhe	575
Qy	1918	ATAAGTAAATAGTATGAACCAAGAAATCCACACTGTGTCTCCCTTTTCAAGCATCAAGT	1977
Dd	576	ValSerHisTyrSerSerValSerThrProCysValHisValTyrLysLeuSerGly	595
Qy	1978	CCTGAAGATGACCAACTCTCAAAACAAAGAAATTTGGCGCACCATTTGGATTCAGCA	2037
Dd	596	ProAspAspAspProLeuHisLysGlnProArgPheTrpAlaSerMetMetGluAlaAla	615
Qy	2038	GGTCTCTTCTGACTATCTCTCCAGAAATTTCTCTTTTGAAGTACTACTCGATT	2097
Dd	616	SerCysProProAspTyrValProProGluLeuPheHisPheHisThrArgSerAspVal	635
Qy	2098	ACATTGATGGGATCTCTACAAGCCTCATGATCTACAGCTCGAAGAAATATCTTACT	2157
Dd	636	ArgLeuTyrGlyMetIleTyrLysProHisAlaLeuGlnProGlyLysLysHisProThr	655
Qy	2158	GTGCTGTTCATATATGGTGCTCTCAGTGCAGTTGGTGAATATCGGTTTAAAGAGTC	2217
Dd	656	ValLeuPheValTyrGlyGlyProGlnValGlnLeuValAenAenSerPheLysGlyIle	675
Qy	2218	AAGTATTTCCGCTTGAATACCTAGCCCTCTAGGTTATGTGGTGTGTAGTATAGACAAAC	2277
Dd	676	LysTyrLeuArgLeuAenThrLeuAlaSerLeuGlyTyrAlaValValValIleAspGly	695
Qy	2278	AGGGATCTCTGACCCAGGCTTAAATTTGAAGGCGCTTTAAATATAAAATGGGTCAA	2337
Dd	696	ArgGlySerCysGlnArgGlyLeuArgPheGluGlyAlaLeuLysAenGlnMetGlyGln	715
Qy	2338	ATAGAATTTGACGATCAGGTGAAGGACTCCAATATCTAGCTTCTCGATATGATTTCAAT	2397
Dd	716	ValGluIleGluAenGlnValGluGlyLeuGlnPheValAlaGluLysTyrGlyPheIle	735
Qy	2398	GACTTAGATCTGTGGGCATCCAGGCTGGTCTTATGGAGGATACCTCTCCCTGATGGCA	2457
Dd	736	AspLeuSerArgValAlaIleHisGlyTrpSerTyrGlyGlyPheLeuSerLeuMetGly	755
Qy	2458	TTAATGCAGAGGTCCAGATATCTTCAGGGTTGCTATTGCTGGGCCCCAGTCACTCTGTGG	2517
Dd	756	LeuIleHisLysProGlnValPheLysValAlaIleAlaGlyAlaProValThrValTyr	775
Qy	2518	ATCTTCTATGATACAGATACACGGAAGCTTATATGGTCCACCTGCACAGAAATGAACAG	2577
Dd	776	MetAlaTyrAspThrGlyTyrThrGluArgTyrMetAspValProGluAenAenGlnHis	795
Qy	2578	GGCTAATTAATAGATCTGTGGCCATCAACGAGAAAGTTCCTCTGACCAACCAATCCT	2637
Dd	796	GlyTyrGluAlaGlySerValAlaLeuHisValGluLysLeuProAenGluProAenArg	815
Qy	2638	TTACTGCTCTTACATGGTTTCTCGATGAGAATGTCCATTTTGCACATACCATATATTA	2697
Dd	816	LeuLeuIleLeuHisGlyPheLeuAspGluAenValHisPheHisThrAenPheLeu	835
Qy	2698	CTGAGTGTTTTGTAGTGGGCTGGAAGCCATATGATTTACAGATCTATCTCTCAGGAGAGA	2757
Dd	836	ValSerGlnLeuIleArgAlaGlyLysProTyrGlnLeuGlnIleTyrProAenGluArg	855
Qy	2758	CACAGCATAAGAGTTCTCTGAATCGGAGAACATTATCAACTGTCATCTTTTGCACCTACCTT	2817
Dd	856	HisSerIleArgCysProGluSerGlyGluHisTyrGluValThrLeuLeuHisPheLeu	875
Qy	2818	CAAGAAAACCTT	2829
Dd	876	GlnGluTyrLeu	879
RESULT 12			
US-10-070-464-5			
; Sequence No. Application US/10070464			
; Patent No. 6881564			
; GENERAL INFORMATION:			
; APPLICANT: GORRELL, Mark Douglas			
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES			
; FILE REFERENCE: GH-007			
; CURRENT APPLICATION NUMBER: US/10/070.464			
; PRIOR FILING DATE: 2002-03-07			
; PRIOR APPLICATION NUMBER: PCT/AU00/01085			
; PRIOR FILING DATE: 2000-09-11			
; PRIOR APPLICATION NUMBER: AU PQ5709			
; PRIOR FILING DATE: 2000-02-18			
; PRIOR APPLICATION NUMBER: AU PQ2762			
; PRIOR FILING DATE: 1999-09-10			
; NUMBER OF SEQ ID NOS: 8			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 5			
; TYPE: PRT			
; ORGANISM: Homo Sapiens			
US-10-070-464-5			
Alignment Scores:			
Pred. No.:	3,528-248	Length:	465
Score:	2422.00	Matches:	465
Percent Similarity:	82.3%	Conservative:	0
Best Local Similarity:	82.3%	Mismatches:	0
Query Match:	43.6%	Indels:	100
DB:	2	Gaps:	1
US-10-825-632-2 (1-3120) x US-10-070-464-5 (1-465)			
Qy	1165	ACAGGTACAGCAAAATCCTAAAGTCATCTTTTAAGATGTCAGAAATATGATGCTGAA	1224
Dd	1	ThrGlyThrAlaAenProLysValThrPheLysMetSerGluIleMetIleAspAlaGlu	20
Qy	1225	GGAAGGATCATAGATGTCTATAGATAAGGAACCTAAATTCACCTTTTGAGATTTCTATTGAA	1284
Dd	21	GlyArgIleIleAspValIleAspLysGluLeuIleGlnProPheGluLeuLeuPheGlu	40
Qy	1285	GGAGTTGAATATATGCCAGAGCTGGATGGACTCCTCCTAGGGGAAATATGCTTGGTCCATC	1344

Db 41 GlyValGluTyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIle 60
QY 1345 CTACTAGATCGCTCCAGACGCTGCTACAGATAGTGTGATCTACACCTGAATATTATC 1404
Db 61 LeuLeuAspArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIle 80
QY 1405 CCAGTAGAAGATGATGTTATGGAAGGACAGAGACTCATTGAGTCAGTGCCTGATCTGTG 1464
Db 81 ProValGluAspValMetGluArgGlnArgLeuIleGluSerValProAspSerVal 100
QY 1465 AGCCCACTAATATTCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTTT 1524
Db 101 ThrProLeuIleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePhe 120
QY 1525 CATGTTTTTCCCAAGTCACGAAGAGAAATGAGTTTATTTTTCCTCTGAGTGCAAA 1584
Db 121 HisValPheProGlnSerHisGluGluIleGluPheIlePheAlaSerGluCysIys 140
QY 1585 ACAGTGTTCCTGATATATACAAATACATCTATTTTAAAGGAAGCAATATAACGA 1644
Db 141 ThrGlyPheArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArg 160
QY 1645 TCCAGTGTGGCTGCTGCTCCCAAGTGATTTCAAGTGTCTCTATCAAGAGGAGTAGCA 1704
Db 161 SerSerGlyGlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAla 180
QY 1705 ATTACACGTGTGAATGGAGATCTTTGGCGGAGTGGATCTAATATCCAAAGTTGATGAA 1764
Db 181 IleThrSerGlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGlu 200
QY 1765 GTCAGAAGCTGGTATATTTTAAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTAC 1824
Db 201 ValArgArgLeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyr 220
QY 1825 GTAGTCAGTTACGTAATCTCTGGAGAGGTGACAAAGGCTGACTACCGTGGCTACTCACAT 1884
Db 221 ValValSerTyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHis 240
QY 1885 TCTTCTGCTGATCAGTCAGTGTGACTTCTTTAAGTAAGTATAGTATACCAAGAAAT 1944
Db 241 SerCysValIleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsn 260
QY 1945 CCACACTGTGTCTCTTTACAGCTATCAAGTCTCAAGTCAAGTCAACCACTGCAAAACA 2004
Db 261 ProHisCysValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThr 280
QY 2005 AAGGAATTTGGGCCACCATTTTGGATTCAGCAGTCTCTCTCTGACTATATCTCTCCA 2064
Db 281 LysGluPheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProPro 300
QY 2065 GAAATTTCTCTTTTGAAGTACTGATTTACATTTGATGATGATGCTCTACAAGCT 2124
Db 301 GluIlePheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysPro 320
QY 2125 CATGATCTACAGCTCGAAAGAAATATCTACTGTGCTGTGTTTATATATATGTTGCTCAG 2184
Db 321 HisAspLeuGlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGln 340
QY 2185 GTGCAGTTGGTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTTAGCC 2244
Db 340 ----- 340
QY 2245 TCTCTAGGTTATGTGTTGTAGTATAGAACAAAGGGGATCTGTCAACGAGGGCTTAAA 2304
Db 340 ----- 340
QY 2305 TTTGAAGGCGCTTTAAATATATAAATGGGTCAAAATAGAAATTGACGATCAGTGGAGGA 2364
Db 340 ----- 340
QY 2365 CTCCATATCTAGCTTCTCGATATGATTTTCAATTGACTTAGATCGTGTGGGCATCCACGGC 2424
Db 340 ----- 340

QY 2425 TGTCTCTATGAGGATACCTCTCCCTGATGGCAATTAATGCAGAGGTGAGATATCTTCAGG 2484
Db 340 ----- 340
QY 2485 GTTGTCTATTGCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGAA 2544
Db 341 ValAlaIleAlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGlu 360
QY 2545 CGTTATATGGTCAACCTGAGCAGCAATGAACAGGCTATTACTTAGGATCTGTGGCCATG 2604
Db 361 ArgTyrMetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMet 380
QY 2605 CAAGCAGAAAGCTTCCCTCTGAACCAATGTTTACTGCTCTTACATGTTTCTCTCGAT 2664
Db 381 GlnAlaGlyLysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAsp 400
QY 2665 GAGAATGTCCATTTTGCACATACCATGATATTTACTGAGTCTTTTAGTCAGGCTGGAAG 2724
Db 401 GluAsnValHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLys 420
QY 2725 CCATATGATTTACAGATCTATCTCTCAGGAGACACACATTAAGAGTTTCTTGAATCGGA 2784
Db 421 ProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGly 440
QY 2785 GAACATTATGAATGTCATCTTTTGCACCTACCTTCAAGAAAACCTTGGATCAGTATTGCT 2844
Db 441 GluHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAla 460
QY 2845 GCTCTAAAAGTGATA 2859
Db 461 AlaLeuLysValIle 465

RESULT 13

US-09-976-674-29
; Sequence 29, Application US/09976674
; Patent No. 6844180

GENERAL INFORMATION:

; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976, 674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 29
; LENGTH: 832
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-29

Alignment Scores:
Pred. No.: 2,07e-246 Length: 832
Score: 2407.00 Matches: 446
Percent Similarity: 71.5% Conservative: 125
Best Local Similarity: 55.8% Mismatches: 198
Query Match: 43.4% Indels: 30
DB: 2 Gaps: 5

US-10-825-632-2 (1-3120) x US-09-976-674-29 (1-832)

QY 124 TGGAGGCGGCGCAGCATGAAGCGGCGCGCTCCATAGCGCAGCTCGGACGCTCC 183
Db 17 TrpArgSerPheSerLeuAsnSerGluGlyAlaGluArgMetAlaThrThrGlyThrPro 36
QY 184 GGGCGGCGCGGGAAGGAAAATGCAACATGGCAGCAGCAATGGAACAGAACAGCTG 243
Db 37 ThrAlaAspArgGlyAsp-----AlaAlaAla----- 45

QY	244	GGTGTGAGATATTTGAAACTCGGACTGTGAGGAGAAATATTGAATCACAGGATCGGCCT	303	QY	1321	GAGGAAAAATATGCTTGGTCCATCTACTAGATCGCTCCGAGCTCCCTACAGATAGTG	1380
Db	46	-----ThrAspAspPro	49	Db	389	AspGlyIysIyrAlaIatrpAlaMetPheLeuAspArgProGlnGlnIrpLeuGlnLeuVal	408
QY	304	AAATTGGAGCCTTTTATGTTGAGCGGTATTTCTTGAGTCAGCTTAAAAAGCTGCTTGCC	363	QY	1381	TTGATCTCACCTGGAATTTATTTATCCAGTAGAAGATGATGTATTGAAAAGGCAGACTC	1440
Db	50	AlaAlaArg---PheGlnValGlnLysHisSerTrpAspGlyLeuArgSerIleIleHis	68	Db	409	LeuLeuProProAlaLeuPheIleProSerThrGluAsnGluGlnGlnArgLeuAlaSer	428
QY	364	GATACCAAAAAATATCATCGCTACATGATGGCTTAAGGACACCATGATTTTCATGTTGTG	423	QY	1441	ATTGAGTCAGTCCGTGATTTCTGTGACGCCACTAATTTATCTATGAAGAAAACAACAGACATC	1500
Db	69	GlySerArgIysTySerGlyLeuIleValAsnLysAlaProHisAspPheGlnPheVal	88	Db	429	AlaArgAlaValProArgAsnValGlnProTyValValIyrGluGluValThrAsnVal	448
QY	424	AAGAGGAATGATCCAGATGGACCTCATTCAGACAGAACTATTATCTTGCCTGCTCTGCT	483	QY	1501	TGGATAAATATCCATGACATCTTTTCATGTTTTCCTCCAAAGTCAC---GAAGAGGAAAT	1557
Db	89	GlnIysThrAspGluSerGlyProHisSerHisArgLeuTyTrpLeuGlyMetProTy	108	Db	449	TrpIleAsnValHisAspIlePheTyProPheProGlnSerGluGlyGluAspGluLeu	468
QY	484	GAGAACAGAAATATACACTGTTTATCTGAAATCCCAAACTCATCAATAGACAGCA	543	QY	1558	GAGTTTATTTTGGCTCTGAAATGCAAAACAGGTTTTCGTCATTTATACAAAATTAATCATCT	1617
Db	109	GlySerArgGluAsnSerLeuLeuTySerGluIleProLysValArgLysGluAla	128	Db	469	CysPheLeuArgAlaAsnGluCysLysThrGlyPheCysHisLeuTyLysValThrAla	488
QY	544	GTCTTAATGCTCTCTGGAAGCCTCTTTGGATCTTTTTCAGGCAACTGGACTATGGA	603	QY	1618	ATTTTAAAGAAAGCAAAATATAACGATCCAGTGGTGGCTGCTGCTCCAAAGTGATTTTC	1677
Db	129	LeuLeuLeuLeuSerTrpLysGlnMetLeuAspHisPheGlnAlaIatrpProHisIleGly	148	Db	489	ValLeuLysSerGlnGlyTyAspTrpSerGluProPheSerProGlyGluAspGluPhe	508
QY	604	ATGATTTCTCGAAGAAGAACTATTAAAGAAAGAAAAACGCAATGGACAGTCGGAAT	663	QY	1678	AAATGCTCTCTATCAAAAGAGGAGATAGCAATACAGTGGTGAATGGGAAGTTCTTGCCCGG	1737
Db	149	ValTySerArgGluGluLeuLeuArgGluArgLysArgLeuGlyValPheGlyIle	168	Db	509	LysCysProIleLysGluGluIleAlaLeuThrSerGlyGluTrpGluValLeuAlaArg	528
QY	664	GCTTCTTACGATTATCACCAAGGAAGTGAACATTTCTGTTTCAAGCCGGTAGTGAAT	723	QY	1738	CATGATCTAATATCCAAAGTTGATGAAGTCAGAGGCTGCTATATTTTGAAGGCACCAAA	1797
Db	169	ThrSerTyAspPheHisSerGluSerGlyLeuPhePheGlnAlaSerAsnSerLeu	188	Db	529	HisGlySerLysIleTrpValAsnGluGluThrLysLeuValTyPheGlnGlyThrLys	548
QY	724	TATCACGTAAGAATGAGGCGCCACAGGATTTTACGCAACCACTTTAAGGCCCAATCTA	783	QY	1798	GACTCCCTTTAGACATCACTCTGATCTGTCAGTCTGTCAGTCACTGAGCTGTGATCTTT	1917
Db	189	PheHisCysArgAspGlyGlyLysAsnGlyPheMetValSerProMetLysProLeuGlu	208	Db	549	AspThrProLeuGluHisLeuTyValValSerTyGluAlaGlyGluIleVal	568
QY	784	GTGGAACACTAGTTGTCCCAACATACGATGATGATCCAAAATTTATGCCCGCTGATCCAGAC	843	QY	1858	AGGCTGACTGACCGGTGCTACTCATCTTCTGTCATCATGCTGAGCTGTGATCTTT	1917
Db	209	IleIysThrGlnCysSerGlyProArgMetAspProLysIleCysProAlaAspProAla	228	Db	569	ArgLeuThrThrProGlyPheSerHisSerCysSerMetSerGlnAsnPheAspMetPhe	588
QY	844	TGGATTGCTTTTATACATAGCAACGATTTTGGATATCTAAACATCTTAACACAGAGAAGAA	903	QY	1918	ATAAGTAAGTATAGTAACCAAGAAATCCACACTGTGTGTCCTTTTACAAGCTATCATAGT	1977
Db	229	PhePheSerPheIleAsnAsnSerAspLeuTrpValAlaAsnIleGluThrGlyGluGlu	248	Db	589	ValSerHisTySerSerValSerThrProProCysValHisValTyLysLeuSerGly	608
QY	904	AGGAGACTCACTTATGTGCACATGAGCTAGCCCAACATGGAAGATGCCAGATCAGCT	963	QY	1978	CCTGAAGATGACCCCAACTTGCAAAACAAAGGAATTTTGGGCCCACTTTTGGATTCAGCA	2037
Db	249	ArgArgLeuThrPheCysHisGlnGlyLeuSerAsnValLeuAspAspProLysSerAla	268	Db	609	ProAspAspAspProLeuHisLysGlnProArgPheTrpAlaSerMetMetGluAlaAla	628
QY	964	GGAGTCGCTACTTTGTTCTCCAAGAAATTTGATAGATATCTCGCTATTTGGTGTGT	1023	QY	2038	GGTCTCTTCTCTGACTATCTCTCCAGAAAATTTTCTTTTGAAGAACTACTCTGGATTT	2097
Db	269	GlyValAlaIatrpPheValIleGlnGluGluPheAspArgPheThrGlyTyTrpIrpCys	288	Db	629	SerCysProProAspTyValProProGluIlePheHisPheHisThrArgSerAspVal	648
QY	1024	CCAAAAGCTGAACAACACTCCAGTGGTGGT---AAAATCTTAGAATCTTATATGAAGAA	1080	QY	2098	ACATTTGATGGGATCTCTACAAGGCTCATCATCTACAGCTCGGAAAGAAATATCTCTACT	2157
Db	289	ProThrAlaSerTrpGluGlySerGluGlyLysThrLeuArgIleLeuTyTrpGluGlu	308	Db	649	ArgLeuTyGlyMetIleTyLysProHisAlaLeuGlnProGlyLysLysHisProThr	668
QY	1081	AATGATGAATCTGAGGTGGAATTTATTCATGTTTACATCCCTATGTTGGAACAAGGAGG	1140	QY	2158	GTGCTGTTTCATATATGCTGCTCCTCAGGTGAGTGGTGGTGAATATCGTTTAAAGGATC	2217
Db	309	ValAspGluSerGluValGluValIleHisValProSerProAlaLeuGluGluArgLys	328	Db	669	ValLeuPheValTyGlyGlyProGlnValGlnLeuValAsnSerPheLysGlyIle	688
QY	1141	GCAGATTCAATTCGCTTATCTTAAACACAGGTACAGCAAACTCTTAAAGCTACTTTTAAAGATG	1200	QY	2218	AGTATTTCCGCTTGAATACCTAGCTCTCTAGGTTATGCTGGTGTAGTATAGACAAAC	2277
Db	329	ThrAspSerTyArgTyProArgThrGlySerLysAsnProLysIleAlaLeuLysLeu	348	Db	689	LysTyLeuArgLeuAsnThrLeuAlaSerLeuGlyTyAlaValValIleAspGly	708
QY	1201	TCAGAAATAATGATGCTGGAAGGAGGATCATAGTCTCATAGATAAGGAACATAAT	1260	QY	2278	AGGGATCTCTGACCGAGGCTTAAATTTGAAGCGCTTTAAATATATAAATCGGTCAA	2337
Db	349	AlaGluPheGlnThrAspSerGlnGlyLysIleValSerThrGlnGluLysGluLeuVal	368	Db	709	ArgGlySerCysGlnArgGlyLeuArgPheGluGlyAlaLeuLysAsnGlnMetGlyGln	728
QY	1261	CAACTTTTGGATCTTATTTGAAGGAGTTGMAATATATTCAGAGCTGGATGGACTCCT	1320	QY	2338	ATAGAAATTCAGCATCAGGTGGAAGGACTCCAAATATCTAGCTTCTCGATATGATTCATT	2397
Db	369	GlnProPheSerSerLeuPheProLysValGluTyIleAlaArgAlaGlyTrpThrArg	388	Db	729	ValGluIleGluAspGlnValGluGlyLeuGlnPheValAlaGluLysTyGlyPheIle	748
QY				QY	2398	GACTTAGATCGTGTGGGCATCCACGGCTGGTCCCTATGAGGATACCTCTCCTCGATGGCA	2457


```
Db 749 AspLeuSerArgValAlaLeHisGlyTrpSerTyrglyGlyPheLeuSerLeuMetGly 768
Qy 2458 TTAATGCGAGGTCAGATATCTTCAGGGTGTCTATTGCTGGGGCCCGACGTCCTGTG 2514
Db 769 LeuIleHisLysProGlnValPheLysAlaGlnProLeuAlaTyProProArgLeu 787

RESULT 14
US-09-976-674-31
; Sequence 31, Application US/09976674
; Patent No. 6844180
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 31
; LENGTH: 832
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-31

Alignment Scores:
Pred. No.: 2,07e-246 Length: 832
Score: 2407.00 Matches: 446
Percent Similarity: 71.5% Conservative: 125
Best Local Similarity: 55.8% Mismatches: 198
Query Match: 43.4% Indels: 30
Db: 2 Gaps: 5

US-10-825-632-2 (1-3120) x US-09-976-674-31 (1-832)
Qy 124 TGGAGGCGCGCAGCATGAAGCGCGCAGCGCGCTCCATAGCGCAGCTGGGAGCGTCC 183
Db 17 TrpArgSerPheSerLeuAenSerGluGlyAlaGluArgMetAlaThrThrGlyThrPro 36
Qy 184 GGGCGGCGCGGGGGAAGGAAAAATGCAACATGGCAGCAGCAGCAATGGAAACAGACGCTG 243
Db 37 ThrAlaAspArgGlyAsp-----AlaAlaAla----- 45
Qy 244 GGTGTTGAGATATTTGAAACTGCGGACTGTGAGGAGAAATATTGAAATCAGAGTCGGCCT 303
Db 46 -----ThrAspAspPro 49
Qy 304 AAATTGGAGCCTTTTATGTTGAGCGGTATCTCTGAGTCAGCTTAAAGCTGTTGCC 363
Db 50 AlaAlaArg---PheGlnValGlnLysHisSerTrpAspGlyLeuArgSerIleIleHis 68
Qy 364 GATACCGAAGAAATATCATGCTACATGATGGCTTAAGCCACCATGATTTCTATGTTGTG 423
Db 69 GlySerArgLysTySerGlyLeuIleValAenLysAlaProHisAspPheGlnPheVal 88
Qy 424 AAGAGGAATGATCAGATGCACTCTTACAGACAGAAATCTATTACCTTGCATGTCGTGT 483
Db 89 GlnLysThrAspGluSerGlyProHisSerHisArgLeuTyTyLeuGlyWepTyPro 108
Qy 484 GAGAACGAGAGAAATACACTGTTTATTCTGAAATTCCTCAAAATCCCAAACTCATATGAGCAGCA 543
Db 109 GlySerArgGluAenSerLeuLeuTySerGluIleProLysValArgLysGluAla 128
Qy 544 GTCCTAATGCTCTCTTGGAGAGCCTCTTTTGGATCTTTTTCAGGCAACACTGACGATATGGA 603
Db 129 LeuLeuLeuLeuSerTrpLysGlnMetLeuAspHisPheGlnAlaThrProHisHisGly 148
Qy 604 ATGTATTCTCGAAGAAGAAGAACTATTAAAGAGAAAGAAACCGCATTTGGACAGCTCGGAAT 663
```

```
Db 149 ValTySerArgGluGluGluLeuLeuArgGluArgLysArgLeuGlyValPheGlyIle 168
Qy 664 GCTTCTTACGATTATCACCAGGAAGTGGACATTTCTGTTTTCAGCGCGGTGGGAATT 723
Db 169 ThrSerTyAspPheHisSerGluSerGlyLeuPheLeuPheGlnAlaSerAsnSerLeu 188
Qy 724 TATCAGCTAAAGATGGAGGGCCACAAAGGATTTACGCAACACCTTTTAAAGCCCAATCTA 783
Db 189 PheHisCysArgAspGlyGlyLysAenGlyPheMetValSerProMetLysProLeuGlu 208
Qy 784 GTGGAATAAGTTGTCCTCCAAACATACGATGGATGCAAAATTTATGCCCGCTGATCCAGAC 843
Db 209 IleLysThrGlnCysSerGlyProArgMetAspProLysIleCysProAlaAspProAla 228
Qy 844 TGGATTCTTTTATACATAGCAACGATATTTGGATATCTAACATCGTAAACAGAGAGAA 903
Db 229 PhePheSerPheIleAenAenSerAspLeuTrpValAlaAenIleGluThrGlyGluGlu 248
Qy 904 AGGAGACTCATTTATGTGCACATGAGCTAGCCACATGGAAGAGAGATGCCAGATCAGCT 963
Db 249 ArgArgLeuThrPheCysHisGlnGlyLeuSerAenValLeuAspAspProLysSerAla 268
Qy 964 GGACTCGCTACCTTTGTTCTCCAAAGAGAAATTTGATAGATATTTCTGGCTATTGGTGGTGT 1023
Db 269 GlyValAlaThrPheValIleGlnGluGluPheAspArgPheThrGlyTyTrpTrpCys 288
Qy 1024 CCAAAAGCTGAAACAACTCCAGTGTGTGT---AAAATCTTTAGAAATCTATATGAAGAA 1080
Db 289 ProThrAlaSerTrpGluGlySerGluGlyLeuLysThrLeuArgIleLeuTyrgluGlu 308
Qy 1081 AATGATGAATCTGAGTGGAAATTTATTCATGTTTACATCCCTTATCTTGGAAACAGAGG 1140
Db 309 ValAspGluSerGluValGluValIleHisValProSerProAlaLeuGluArgLys 328
Qy 1141 GCAGATTCAATTCGGTTATCTCTAAACAGGTACAGCAATCTTAAAGTCACATTTTAAATG 1200
Db 329 ThrAspSerTyArgTyProArgThrGlySerLysAenProLysIleAlaLeuLysLeu 348
Qy 1201 TCAGAAATATGATGATGCTGAAGAGATCATAGATGTCATAGATAAGAACTAAT 1260
Db 349 AlaGluPheGlnThrAspSerGlnGlyLysIleValSerThrGlnGlyLysGluVal 368
Qy 1261 CAACCTTTGAGATTCCTATTGAGAGTGTGAATATATTCAGAGCTGGATGGAGCTCCT 1320
Db 369 GlnProPheSerSerLeuPheProLysValGluTyIleAlaArgAlaGlyTrpThrArg 388
Qy 1321 GAGGAAATATGCTGCTCCATCTCTAGATCGCTCCAGACTCCGCTACAGATAGTG 1380
Db 389 AspGlyLysTyAlaTrpAlaMetPheLeuAspArgProGlnGlnTrpLeuGlnLeuVal 408
Qy 1381 TTGATCTCACCTGAATTTATTTATCCAGTAGAAGATGATGTTATGGAAGGAGCAGACTC 1440
Db 409 LeuLeuProProAlaLeuPheIleProSerThrGluAenGluGluGlnArgLeuAlaSer 428
Qy 1441 ATTGAGTCAGTGCCTGATCTCTGACGCGCACTAATTTATCTATGAAGAAACACAGACATC 1500
Db 429 AlaArgAlaValProArgAenValGlnProTyTyValValTyGluGluValThrAenVal 448
Qy 1501 TGGATAATATCCATGACATCTTTTCATGTTTTCCTCCCAAGTCAC---GAAAGGAAAT 1557
Db 449 TrpIleAenValHisAspIlePheTyProPheProGlnSerGluGlyGluAspGluLeu 468
Qy 1558 GAGTTATTTTGGCTCTGATGCAAAACAGAGTTTCCTGTCATTTATACAAAATTTACATCT 1617
Db 469 CysPheLeuArgAlaAenGluCysLysThrGlyPheCysHisLeuTyrgLysValThrAla 488
Qy 1618 ATTTTAAAGGAAAGCAAAATATAAACGATCCAGTGGTGGCTGCTCTCCAAAGTGAATTC 1677
Db 489 ValLeuLysSerGlnGlyTyArgTrpSerGluProPheSerProGlyGluAspGluPhe 508
Qy 1678 AAGTGTCTCTATCAAGAGGAGATAGCAATTTACCAGTGGTGAATGGGAAGTTCTTGGCCGG 1737
```


509	LysCysProIleLysGluGluIleAlaLeuThrSerGlyGluTrpGluValLeuAlaArg	528
1738	CATGGATCTAATATCCAAAGTTGATGAAGTCAGAAAGCGCTGGTATATTTTGAAGGACCAAA	1797
529	HisGlySerLysIleTrpValAenGluThrLysLeuValTyPheGlnGlyThrLys	548
1798	GACTCCCTTTAGAGCATCCCTGTACGTAGTCAGTTACGTAAATCCCTGGAGAGGTGACA	1857
549	AspThrProLeuGluHisLeuValValValValValValValValValValValValValVal	568
1858	AGGCTGACTGACCGTGCTACTCACATTTCTGCTGATCAGTCAGTCAGTCAGTCAGTCAGTC	1917
569	ArgLeuThrProGlyPheSerHisSerCysSerMetSerGlnAenPheAspMetPhe	588
1918	ATAAGTAAATAGTATACCAAGAGAAATCCACATCTGCTCTTACAGCTATCAAGT	1977
589	ValSerHisTySerSerValSerThrProProCysValHisValTyLysLeuSerGly	608
1978	CCTGAAGATGACCCAACTTGCANAAACAAAGGAATTTGGCCACCATTTTGGATTTCAGCA	2037
609	ProAspAspAspProLeuHisLysGlnProArgPheTrpAlaSerMetMetGluAlaAla	628
2038	GGTCTCTTCTGACTATPACTCTCCAGAAATTTCTCTTTTGAAGTACTACTCGATT	2097
629	SerCysProProAspTyValProProGluIlePheHisPheHisThrArgSerAspVal	648
2098	ACATTTGTATGGGATGCTCTACAAGCCTCATGATCTACAGCCTGGAAAGAAATATCTACT	2157
649	ArgLeuTyGlyMetIleTyLysProHisAlaLeuGlnProGlyLysLysHisProThr	668
2158	GTGCTGTTCAATATATGGTGTCTCCTCAGGTGCAGTTGGTGAATATCGGTTTAAAGAGTC	2217
669	ValLeuPheValTyGlyGlyProGlnValGlnLeuValAenAenSerPheLysGlyIle	688
2218	AAGTATTTCCGCTTCAATACCTTAGCCTCTAGGTATGCTGTTAGTGTAGTAGACAAAC	2277
689	LysTyLeuArgLeuAenThrLeuAlaSerLeuGlyTyAlaValValValIleAspGly	708
2278	AGGGATCTGTCCACCGAGGCTTAAATTTGAAGGCGCTTTAAATATAAATATGGTCAA	2337
709	ArgGlySerCysGlnArgGlyLeuArgPheGluGlyAlaLeuLysAenGlnMetGlyGln	728
2338	ATAGAANTGACATCAGTCAGTGAAGGACTCCCATATCTAGCTTCTCGATATGATTTCAAT	2397
729	ValGluIleGluAspGlnValGluGlyLeuGlnPheValAlaGluLysTyGlyPheIle	748
2398	GACTTAGATCGTGGGATCCAGGCTGTGCTCTATCGAGTACCTCTCCCTGATGGCA	2457
749	AspLeuSerArgValAlaIleHisGlyTrpSerTyGlyGlyPheLeuSerLeuMetGly	768
2458	TTAATCGAGGTGACATATCTTCAGGTTCTGATTCGTTGCTGGGCGCCAGTCACCTCTG	2514
769	LeuIleHisLysProGlnValPheLysAlaGlnProLeuAlaTyProProArgLeu	787
RESULT 15		
US-09-976-674-37		
; Sequence 37, Application US/09976674		
; Patent No. 6844180		
; GENERAL INFORMATION:		
; APPLICANT: Qi, Steve		
; APPLICANT: Akinsanya, Karen		
; APPLICANT: Riviere, Pierre		
; APPLICANT: Junien, Jean-Louis		
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPP1V		
; FILE REFERENCE: 70669		
; CURRENT APPLICATION NUMBER: US/09/976,674		
; PRIOR FILING DATE: 2001-10-12		
; PRIOR APPLICATION NUMBER: US 60/240,117		
; PRIOR FILING DATE: 2000-10-12		
; NUMBER OF SEQ ID NOS: 61		
; SOFTWARE: Patentin version 3.1		
; SEQ ID NO 37		
; LENGTH: 819		
; TYPE: PRT		
; ORGANISM: Homo sapiens		
US-09-976-674-37		
Alignment Scores:	3.8e-241	Length: 819
Pred. No.:	2357.50	Matches: 439
Score:	70.3%	Conservative: 123
Percent Similarity:	54.9%	Mismatches: 134
Best Local Similarity:	42.5%	Indels: 43
Query Match:	2	Gaps: 6
DB:		
US-10-825-632-2 (1-3120) x US-09-976-674-37 (1-819)		
QY	124 TGGAGCGCGCGCAGCATGAAAGCGCGCGCGCTCCATAGCGCAGCTCGGAGCGGTCC	183
DB	17 TrpArgSerPheSerLeuAenSerGluGlyAlaGluArgMetAlaThrThrGlyThrPro	36
QY	184 GGGCGGGCGCGGGGGAAGGAAAATGCAACATCGCAGCAGCAATGGAAAACAGACAGCTG	243
DB	37 ThrAlaAspArgGlyAsp-----AlaAlaAla-----	45
QY	244 GGTGTTGAGATATTTTGAAACTGCGGACTGTGAGGAGAAATATTGAATCACAGGATCGGCT	303
DB	46 -----	49
QY	304 AAATTGGAGCCTTTTATGTTGAGCGGTATTCTCGGAGTCAGCTTAAAGAGCTGTGCC	363
DB	50 AlaAlaArg---PheGlnValGlnLysHisSerTrpAspGlyLeuArgSerIleIleHis	68
QY	364 GATACACAGAAAATATCATGCTACATGCTAAGCAGCACCATCATGATTTTCATGTTGTG	423
DB	69 GlySerArgGlyTySerGlyLeuIleValAenLysAlaProHisAspPheGlnPheVal	88
QY	424 AAGAGGAATGATCCAGATGAGCAGCTTACAGACAGATCTATTACCTTGCATGCTCGT	483
DB	89 GlnLysThrAspGluSerGlyProHisSerHisArgLeuLeuTyTrpLeuGlyMetProTy	108
QY	484 GAGAACAGAGAAAATACATGTTTATCTGAATTCCTCAAAATCCAAATCATCAATGAGCA	543
DB	109 GlySerArgGluAenSerLeuLeuTySerGluIleProLysLysValArgLysGluAla	128
QY	544 GTCTTAATGCTCTCTGTGAAGCCTCTTTTGGATCTTTTCAGCGACACCTGGACTGGA	603
DB	129 LeuLeuLeuLeuSerTrpLysGlnMetLeuAspHisPheGlnAlaThrProHisHisGly	148
QY	604 ATGATATCTCGAGAGAGAACTATTAAAGAGAAAAGAAACGATTTGGAACAGTCGGAAAT	663
DB	149 ValTySerArgGluGluLeuLeuArgGluArgLysArgLeuGlyValPheGlyIle	168
QY	664 GCTTCTTACGATTTATCACCAGGAAGTGAACATTTCTGTTTCAAGCCGCTAGTGAAT	723
DB	169 ThrSerTyAspPheHisSerGluSerGlyLeuPheLeuPheGlnAlaSerAenSerLeu	188
QY	724 TATCAGTAAAAAGATGGAGGCGCACAGGATTTACGCAACCTTTTAAGCGCCCAATCTA	783
DB	189 PheHisCysArgAspGlyGlyLysAenGlyPheMetValSerProMetLysProLeuGlu	208
QY	784 GTGGAACCTAGTTGTCCCAACATACGATGGATCCAAATTTATGCCCGCTGATCCAGAC	843
DB	209 IleLysThrGlnCysSerGlyProArgMetAspProLysIleCysProAlaAspProAla	228
QY	844 TGGATTCCTTTATACATAGCAACGATATTGGATATCTTAACATCGTAAACAGAGAGAA	903
DB	229 PhePheSerPheIleAenAenSerAspLeuTrpValAlaAenIleGluThrGlyGluGlu	248
QY	904 AGGAGACTCATTTATGTGCACAAATGAGCTAGCCACATGGAGAGATGCCAGATCAGCT	963
DB	249 ArgArgLeuThrPheCysHisGlnGlyLeuSerAenValLeuAenAspProLysSerAla	268
QY	964 GGAGTCGCTACCTTTGTTCTCCAGAGAAATTTAGATATTCTGCTGCTATTGCTGCTGT	1023
DB	269 GlyValAlaThrPheValIleGlnGluGluPheAspArgPheThrGlyTyTrpTrpCys	288

QY	1024	CCAAAGCTGAAACAACTCCCACTGGTGGT---AAAATCTTAGAATTTCTATATGAAGAA	1080	QY	2098	ACATTGTATGGATGCTCTACAGCCTCATGATCTACAGCTTGGAAAGAAATATCCTACT	2155
DB	289	ProthrAlaSerTrpGluGlySerGluGlyLeuLysThrLeuArgGileLeuTyrgluGlu	308	DB	636	ArgLeuTyrglyMetIleTyrglyProHisAlaLeuGlnProGlyLysLysHisProThr	655
QY	1081	AATGATCAATCTCAGGTGGAATATTATTCATGTTACATCCCCCTATGTTGGAAACAAGGAGG	1140	QY	2158	GTGCTGTTTCATATATGTTGGTCCCTCAGGTGCAGTTGGTGAATAATCGTTTAAAGGATC	2217
DB	309	ValAspGlnSerGluValGluValIleHisValProSerProAlaLeuGluGluArgLys	328	DB	656	ValLeuPheValTyrglyGlyProGlnValGlnLeuValAsnAsnSerPheLysGlyIle	675
QY	1141	GCAGATTCAATTCGGTTATCTCTAAAACAGGTACAGCAAATCTCTAAAGTCACATTTTAAAGATG	1200	QY	2218	AGTATTTCCGCTTGAATACCTCCTAGCTCTCTAGTTATGTGGTTGTAGTGATGACACAAC	2277
DB	329	ThrAspSerTyrglyArgTyrglyProArgThrGlySerLysAsnProLysIleAlaLeuLysLeu	348	DB	676	LysTyrglyLeuArgLeuAsnThrLeuAlaSerLeuGlyTyrglyAlaValValIleAspGly	695
QY	1201	TCGAAATAATGATTGATGCTGAAGGAAGGATCATAGATGTCATAGATAAGCAACTAATT	1260	QY	2278	AGGGATCTGTCCAGGAGGCTTAAATTTCAAGCGCCTTTAAATATAATAATCGGTCAA	2337
DB	349	AlaGluPheGlnThrAspSerGlnGlyLysIleValSerThrGlnGlnLysGluLeuVal	368	DB	696	ArgGlySerCysGlnArgGlyLeuArgPheGluGlyAlaLeuLysAsnGlnMetGlyGln	715
QY	1261	CAACCTTTTCAGATTTCTATTGAAGGAGTTGAATATATTGCCAGAGCTGGATGCACTCCT	1320	QY	2338	ATAGAAATTTGACGATCAGGTGGAGGACTCCAATATCTAGCTTCTCGATATGATTCATT	2397
DB	369	GlnProPheSerSerLeuPheProLysValAluTyriIleAlaArgAlaGlyTrpThrArg	388	DB	716	ValGluIleGluAspGlnValGluGlyLeuGlnPheValAlaGluLysTyrglyPheIle	735
QY	1321	GAGGAAAATATGCTGGTCCATCTCTAGATCGCTCCAGACTCGCTACAGATAGTG	1380	QY	2398	GACTTAGATCGTGTGGGCATCCAGGCTGGTCCCTATGGAGATACCTCTCCCTGATGCCA	2457
DB	389	AspGlyLysTyrglyAlaTrpAlaMetPheLeuAspArgProGlnGlnTrpLeuVal	408	DB	736	AspLeuSerArgValAlaIleHisGlyTrpSerTyrglyGlyPheLeuSerLeuMetGly	755
QY	1381	TTGATCTCACTGAAATTTATTATCCCACTAGAAGATGATGTTTATGGAAGGCAGAGACTC	1440	QY	2458	TTAATGCAGAGGTGAGATATCTTCAGGGTTCCTATTGCTGGGGCCCCCAGTCACTCTG	2514
DB				DB	756	LeuIleHisProGlnValPheLysAlaGlnIleLeuAlaTyrglyProArgLeu	774

Search completed: April 14, 2006, 11:37:46
Job time : 181 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: April 14, 2006, 12:23:10 ; Search time 86 Seconds
(without alignments)
3031.692 Million cell updates/sec

Title: US-10-825-632-2
Perfect score: 5552
Sequence: 1 aagtgctaaagctccgagg.....agaattactaaaaaaa 3120

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1867569 seqs, 417829326 residues
Total number of hits satisfying chosen parameters: 3735138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlsp
-Q=/abss/ABSSWEB spool/US10825632/runat_14042006_092359_3688/app_query.fasta_1
-DB=Published Applications AA Main -QFMT=fastan -SUFFIX=n2p.rapbm
-MTNMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct
-THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptco -NORM=ext
-HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abes02p
-USER=US10825632 @CEN 1 1 307 @runat_14042006_092359_3688 -NCPU=6 -ICPU=3
-NO WMAP -NEG SCORES=0 -WAIT -DSPBLOC=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA Main:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4700	84.7	882	3	US-09-976-674-1
2	4700	84.7	882	4	US-10-054-776-2
3	4700	84.7	882	4	US-10-170-789-38
4	4700	84.7	882	4	US-10-311-035-9
5	4700	84.7	882	4	US-10-072-012-622
6	4700	84.7	882	4	US-10-415-122-6
7	4700	84.7	882	4	US-10-825-632-1
8	4700	84.7	882	5	US-10-982-512-1
9	4528.5	81.6	883	4	US-10-072-012-621
10	3607.5	65.0	690	3	US-09-976-674-7
11	3607.5	65.0	690	5	US-10-982-512-7

12	3513.5	63.3	661	3	US-09-976-674-11	Sequence 11, Appl
13	3513.5	63.3	661	5	US-10-982-512-11	Sequence 11, Appl
14	3504	63.1	658	3	US-09-976-674-19	Sequence 19, Appl
15	3504	63.1	658	5	US-10-982-512-19	Sequence 19, Appl
16	3236	58.3	613	3	US-09-976-674-21	Sequence 21, Appl
17	3236	58.3	613	5	US-10-982-512-21	Sequence 21, Appl
18	2871	51.7	892	3	US-09-976-674-23	Sequence 23, Appl
19	2871	51.7	892	5	US-09-976-674-27	Sequence 27, Appl
20	2871	51.7	892	5	US-10-982-512-27	Sequence 23, Appl
21	2871	51.7	892	5	US-10-982-512-27	Sequence 27, Appl
22	2871	51.7	892	5	US-10-433-757-12	Sequence 12, Appl
23	2871	51.7	892	4	US-10-415-122-2	Sequence 2, Appl
24	2870	51.7	863	3	US-09-976-674-3	Sequence 3, Appl
25	2870	51.7	863	4	US-10-072-012-619	Sequence 619, App
26	2870	51.7	863	5	US-10-982-512-3	Sequence 3, Appl
27	2862	51.5	863	4	US-10-072-012-224	Sequence 224, App
28	2862	51.5	863	4	US-10-072-012-226	Sequence 226, App
29	2852.5	51.4	580	4	US-10-275-505-2	Sequence 2, Appl
30	2852.5	51.4	580	6	US-11-140-224-2	Sequence 2, Appl
31	2835	51.1	830	4	US-10-415-122-7	Sequence 7, Appl
32	2833	51.0	869	4	US-10-415-122-4	Sequence 4, Appl
33	2821.5	50.8	879	3	US-09-976-674-33	Sequence 33, Appl
34	2821.5	50.8	879	3	US-09-976-674-35	Sequence 35, Appl
35	2821.5	50.8	879	5	US-10-982-512-33	Sequence 33, Appl
36	2821.5	50.8	879	5	US-10-982-512-35	Sequence 35, Appl
37	2422	43.6	465	4	US-10-825-632-5	Sequence 5, Appl
38	2407	43.4	832	3	US-09-976-674-29	Sequence 29, Appl
39	2407	43.4	832	3	US-09-976-674-31	Sequence 31, Appl
40	2407	43.4	832	5	US-10-982-512-29	Sequence 29, Appl
41	2407	43.4	832	5	US-10-982-512-31	Sequence 31, Appl
42	2397	43.2	689	4	US-10-072-012-620	Sequence 620, App
43	2357.5	42.5	819	3	US-09-976-674-37	Sequence 37, Appl
44	2357.5	42.5	819	3	US-09-976-674-39	Sequence 39, Appl
45	2357.5	42.5	819	5	US-10-982-512-37	Sequence 37, Appl

ALIGNMENTS

RESULT 1
US-09-976-674-1
; Sequence 1, Application US/09976674
; Patent No. US20020115843A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976,674
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; TYPE: PRT
; LENGTH: 882
; ORGANISM: Homo sapiens
US-09-976-674-1

Alignment Scores:
Pred. No.: 0
Score: 4700.00
Percent Similarity: 100.0%
Best Local Similarity: 100.0%
Query Match: 84.7%
DB: 3
Length: 882
Matches: 882
Conservative: 0
Mismatches: 0
Indels: 0
Gaps: 0

US-10-825-632-2 (1-3120) x US-09-976-674-1 (1-882)

QY 214 ATGCACGACCAATGCAACAGCTGGCTGGATTTTGAATTTTGAACCTGCGACTGT 273
|||||

Db	1	MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys	20
Qy	274	GAGGAGAAATATCAATCACAGGATCGCCCTAAATTTGGAGCCCTTTTATGTTGACGGGTAT	333
Db	21	GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr	40
Qy	334	TCCTGGAGTCAGCTTAAAGCTGCTGGATACGATACAGAAAATATCATGGCTACATGATG	393
Db	41	SerTrpSerGlnLeuIleLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet	60
Qy	394	GCTAAGGACCAATATGATTTTCATGTTTGTGAAGAGATGATCCAGATGGACCTCATCA	453
Db	61	AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProHisSer	80
Qy	454	GACAGAACTATATACCTGTCATGCTGTGTGTGAGAACAGAGAAAATACACTGTTTATCT	513
Db	81	AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer	100
Qy	514	GAATTTCCAAAACATCAATAGACGACGATCTTAATGCTCTCTTGGAGCGCTCTTTTG	573
Db	101	GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu	120
Qy	574	GATCTTTTTCAGGCAACACTGGACTATGGAATGTATTTCTCGAAGAAAGAACTATTAAAG	633
Db	121	AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg	140
Qy	634	GAAGAAAACCATTTGAAACAGTCGGAAATGCTTCTTACGATTTATCACCAAGGAAGTGA	693
Db	141	GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly	160
Qy	694	ACATTTCTGTTTCAAGCGGTAGTGGAAATTTATCACTAAAGATGGAGGGCCACAGGA	753
Db	161	ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly	180
Qy	754	TTTAGCAACACCTTTAAAGCCCAATCTAGTGGAACTAGTTGTCCCAACATACGGATG	813
Db	181	PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet	200
Qy	814	GATCAAAATTTATGCCCGCTGATCCAGATCGATGATGCTTTTATACATAGCAACCATAT	873
Db	201	AspProLysLeuCysProAlaAspProAspIleAlaPheIleHisSerAsnAspIle	220
Qy	874	TGGATATCTAATCGTTAAACAGAGAAAGAGAGACTCATTTATGTGCAATGCTA	933
Db	221	TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu	240
Qy	934	GCCAAATGGAAGAGATGCCAGATCAGCTGGAGTGGTACCTTTGTTCTCCCAAGAGAA	993
Db	241	AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu	260
Qy	994	TTTGATAGATATCTGGCTATTTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1053
Db	261	PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly	280
Qy	1054	AAAAATCTTAGAATTTCTATATGAAGAAAATGATGAATCTGAGGTGGAAAATTTATCATGTT	1113
Db	281	LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal	300
Qy	1114	ACATCCCTCATGTTTGGAAACAGGAGGGCAGATTCAATTCGTTATCTTAAACAGGTACA	1173
Db	301	ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr	320
Qy	1174	GCAATCTTAAGTACATTTTAAAGATGTCAGAAATATGATGATGCTGAAGGAAGATC	1233
Db	321	AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle	340
Qy	1234	ATAGATGTCATAGATAAGAACTAATTCACCTTTTGGATTTCTATTGGAAGAGTTGAA	1293
Db	341	IleAspValIleAspLysGluIleGlnProPheGluIleLeuPheGluGlyValGlu	360
Qy	1294	TATATTGCGAGAGCTGGATGGATCTCTGAGGAAAATATGCTGGTCCATCTCTACTAGAT	1353
Db	361	TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp	380
Qy	1354	CGCTCCAGACTCCCTTACAGATAGTGTGTGATCTCTCACTGAATTTATTTATCCAGTAGAA	1413
Db	381	ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu	400
Qy	1414	GATGATGTTATGGAAGGACAGACTCATTTAGTCTAGTGTCTGATTTCTGTGAGCCACTA	1473
Db	401	AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu	420
Qy	1474	ATTATCTATGAAGAAAACACAGACATCTGATATAATATCATGACATCTTTTCAATGTTTT	1533
Db	421	IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe	440
Qy	1534	CCCCAAGTTCACGAAGAGGAAATAGCTTTATTTTGGCTCTCAATGCAAAAACAGGTTTC	1593
Db	441	ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe	460
Qy	1594	CGTCAATTTATCAAAATTTACATCTATTATAAGGAAGCAAAATATAAAGATCCAGTGGT	1653
Db	461	ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly	480
Qy	1654	GGGCTGCTGCTCCCAAGTGTTCCTCAAGTGTCTTCAAGAGGAGATAGCAATTTACCACT	1713
Db	481	GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer	500
Qy	1714	GGTGAATGGAAAGTCTTGGCCGCATCGATCTAATATCCAAGTTCAGTGAAGTCAGAAG	1773
Db	501	GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArg	520
Qy	1774	CTGGTATATTTGAAGGCCCAAAAGACTCCCTTTTAGAGCATCACCTGTACGTAGTCAGT	1833
Db	521	LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSer	540
Qy	1834	TACGTAATCTCTGGAGAGGTGACAAAGCTGATGACCGTGGCTACTCACATTTCTGCTGC	1893
Db	541	TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCys	560
Qy	1894	ATCAGTCACACACTGTGACTTCTTTATAGTAAGTATAGTAAACAGAAATCCACACTGT	1953
Db	561	IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys	580
Qy	1954	GTGTCCTTTTCAAGCTATCAAGTCTCAAGATCACCACTGTGCCAACTTGCAGAAATTTT	2013
Db	581	ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe	600
Qy	2014	TGGGCCACCATTTTGGATTTCAGCAGGTCTCTCTCTGACTATATCTCTCCAGAAATTTTC	2073
Db	601	TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePhe	620
Qy	2074	TCTTTTGAAGTACTACTGATTTTACATTTGTATGGATGCTCTCAAGCCCTCATGATCTA	2133
Db	621	SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu	640
Qy	2134	CAGCTTGAAGAAATATCTACTGCTGCTGTGTTTCATATATGCTGTGTCTCAGGTGCAGTTG	2193
Db	641	GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu	660
Qy	2194	GTGAATAATCGGTTTAAAGGAGTCAAGTATTTCCGTTTCCGTTTAAATTTGAAGGC	2253
Db	661	ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuHisSerLeuGly	680
Qy	2254	TATGTGTTGTAGTATAGACAAACAGGGGATCTCTGTCTCCGAGGGCTTAAATTTGAAGC	2313
Db	681	TyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheIleGly	700
Qy	2314	GCCTTTTAAATATAAATGGGTCAAAATPAGAAATGACATCAGGTGGGAAGGACTCCATAT	2373
Db	701	AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr	720
Qy	2374	CTAGCTTCTCGATATGATTTTCACTTATGATCTGTGGGGCATCCAGGCTGGTCTTAT	2433
Db	721	LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyr	740

QY 2434 GGAGGATACCTCTCCCTGATGGCAATTAATGACAGAGGTGAGATATCTTCAGGGTGTCTATT 2493
DB 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
QY 2494 GCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGATACACGGAACTTATATG 2553
DB 761 AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780
QY 2554 GGTCAACCTGACAGAGTAAGACAGGGCTATTACTTAGATCTGTGGCCATGCAAGCAGAA 2613
DB 781 GlyHisProAspGlnAsnGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
QY 2614 AGTTCCCTCTGAACCAATCGTTTACTGCTCTTACATGCTTCTCGGATGAGATGTC 2673
DB 801 LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
QY 2674 CATTTTGCACATACAGATATATTACTGAGTTTTTTAGTGAGGGCTGGAAAGCCATATGAT 2733
DB 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
QY 2734 TTACAGATCTATCCTCAGGAGACACAGACATAAGAGTTCTGAAATCGGAGAACATTAT 2793
DB 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
QY 2794 GAACCTGATCTTTTGCACCTTCAAGAAACCTTGGATCACGATTTGCTGCTCTAAAA 2853
DB 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeuLys 880
QY 2854 GTGATA 2859
DB 881 ValIle 882

RESULT 2

US-10-054-776-2
; Sequence 2, Application US/10054776
; Publication No. US20030165818A1
; GENERAL INFORMATION:
; APPLICANT: Mark Robert Edbrooke
; APPLICANT: Alan Peter Lewis
; TITLE OF INVENTION: NOVEL PROTEIN
; FILE REFERENCE: QG1042US
; CURRENT APPLICATION NUMBER: US/10/054,776
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-054-776-2

Alignment Scores:
Pred. No.: 0 Length: 882
Score: 4700.00 Matches: 882
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 84.7% Indels: 0
DB: 4 Gaps: 0

US-10-825-632-2 (1-3120) x US-10-054-776-2 (1-882)

QY 214 ATGGCAGCAGCAATGAAACAGACAGCTGGGTGTGAGATATTTGAACTGGCGACTGT 273
DB 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
QY 274 GAGCAGATATTCGAATCACAGGATCGGCTAAATTTGGAGCCCTTTTATGTTGAGCGGTAT 333
DB 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
QY 334 TCCTGGAGTCAAGCTTAAAGCTTGGCCGATACCAAGAAATATCATGGCTACATGATG 393
DB 41 SerTrpSerGlnLeuLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60

QY 394 GCTAAGCCACCATGATGATTTTCATGTTTCTGAAGAGGAATGATCCAGATGGACCTCATTC 453
DB 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
QY 454 GACAGAACTTATACCTTGCCTGCTGCTGAGAACAGAGAAAATACACTGTTTATTCT 513
DB 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
QY 514 GAATTTCCCAAACTATCAATAGACAGCAGCTCTTAATGCTCTCTCTGGAAGCCCTCTTTTG 573
DB 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
QY 574 GATCTTTTTCAGCAACACTGCACTATGGAATGTATTCTCCAGAGAAGAACTATTAAAGA 633
DB 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuLeuArg 140
QY 634 GAAAGAAAACGCATTGGAACAGTCGGAATTTGCTTCTTACGATTATCACCAAGGAAGTGA 693
DB 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
QY 694 ACATTTCTGTTTCAAGCCGTAGTGAATTTATACGTAAAGATGAGAGGCCCAAGAGA 753
DB 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
QY 754 TTACGCAACAACTTTAAGCCCAATCTAGTGGAACTAGTTGTGCCAACATACGATG 813
DB 181 PheThrGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
QY 814 GATCCAAAATATTGCCCCCTGATCCAGACTGGATGCTTTTATACATAGCAACGATATT 873
DB 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
QY 874 TGGATATCTAAACATGCTTAACACAGAGAGAGAGACTCACTTATGTGCACAATGAGCTA 933
DB 221 TrpIleSerAsnIleValThrArgGluGluArgGluThrTyrValHisGlnLeuLeu 240
QY 934 GCCAATGGAAGAAGATGCCAGATCAGCTGGAGTGCCTACCTTTGTTTCTCCAAAGAGAA 993
DB 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu 260
QY 994 TTTGATAGATATTTCTGGCTATTGCTGTTGCTGCTCAAAAGCTGAAACAACCTCCAGTGGT 1053
DB 261 PheAspArgTyrSerGlyTyrTrpCysProLysAlaGluThrThrProSerGlyGly 280
QY 1054 AAAATTTTAGAATTTCTATATGAAGAAAATGATGAATCTGAGGTGGAAAATTTATTCATGTT 1113
DB 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHisVal 300
QY 1114 ACATCCCTTATGTTGAAAACAAGGAGGGCAGATTCAATTCCTGTTATCTCTAAACAGGTACA 1173
DB 301 ThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThr 320
QY 1174 GCAANTCCTAAAGTCACTTTTAAGATGTCCAGAAAATAATGATGTCTGAAGGAGGATC 1233
DB 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
QY 1234 ATAGATGCTCATAGATAAGGAACCTAAATTCACCTTTTGTAGATCTTATTTGAAGGAGTTGAA 1293
DB 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
QY 1294 TATATTGCCAGAGCTGGATGGACTCCTGAGGGAAAATAATGCTGTGCTCCATCTACTAGAT 1353
DB 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
QY 1354 CGCTCCAGACTCGCTACAGATAGTGTGATCTCACTGAAATTTATTTATCCCAAGTAGAA 1413
DB 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
QY 1414 GATGATGTTATGAAAAGGAGAGACATCATTTAGTGCCTGCTGATTTCTGTGACGCCACTA 1473
DB 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
QY 1474 ATTATCTATCAAGAAACAACACACATCTCGATAAATATCCATGACATCTTTTCATGTTT 1533

;
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: US 09/961,721
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: PCT/US01/29904
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: US 60/235,023
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US 10/045,367
; PRIOR FILING DATE: 2001-11-07
; PRIOR APPLICATION NUMBER: US 60/246,561
; PRIOR FILING DATE: 2000-11-07
; PRIOR APPLICATION NUMBER: US 09/801,275
; PRIOR FILING DATE: 2001-03-06
; PRIOR APPLICATION NUMBER: PCT/US01/07074
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: US 60/187,420
; PRIOR FILING DATE: 2000-03-07
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-170-789-38

Alignment Scores:
Pred. No.: 0 Length: 882
Score: 4700.00 Matches: 882
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 84.7% Indels: 0
DB: 4 Gaps: 0

US-10-825-632-2 (1-3120) x US-10-170-789-38 (1-882)

QY	214	ATGCGACGACAAATGGAACAGAACAGCTGGGTGTGAGATATTTGAAACTGCGACTGT	273
DB	1	MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys	20
QY	274	GAGGAGAAATTTGAATCACAGGATCGGCTTAATTTGGAGCCCTTTTATTTGAGCGGTAT	333
DB	21	GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr	40
QY	334	TCCTGGAGTCAGTTAAAGCTGCTTGGCGATACAGAAATATCATGGCTACATGATG	393
DB	41	SerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet	60
QY	394	GCTAAGCCACCATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTC	453
DB	61	AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer	80
QY	454	GACAGATCTATTACCTTGCATGCTGCTGTGAGAACAGAGAAATACACTCTTTTATTCT	513
DB	81	AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer	100
QY	514	GAAATCCCAAACTATCAATAGACAGCAGCTTAAATGCTCTCTTTGGAAGCCCTTTTGTG	573
DB	101	GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu	120
QY	574	GATCTTTTTCAGGCAACACTGGACTATGGAATGTATTCTCGAGAGAGAAAGAACTATTAGA	633
DB	121	AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg	140
QY	634	GAAGAAACCGCATTTGAACAGTCGGAATTTGCTTCTTACGATTATCACCAAGGAGTGG	693
DB	141	GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly	160
QY	694	ACATTCTGTTTCAAGCCGTAGTGGAAATTTATACGTTAAAGATGGAGGCCCAAGGA	753
DB	161	ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly	180
QY	754	TTTACGCAACAACTTTAAGCCCAATCTAGTGGAAACTAGTTGTGCCCAACATACGGATG	813

DB	181	PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet	200
QY	814	GATCCAAATTTATGCCCGCTGATCCAGACTGGATTGCTTTTATACATAGCAACGATATT	873
DB	201	AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle	220
QY	874	TGGATATCTAACATCGTAAACCAGAGAGAAAGAGACTCACCTTATGTGCACAATAGACTA	933
DB	221	TrpIleSerAsnIleValThrArgGluArgGluLeuThrTyrValHisAsnGluLeu	240
QY	934	GCCAAACATGAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCAAGAAGAA	993
DB	241	AlaAsnMetGluGluAspAlaAspSerAlaGlyValAlaThrPheValLeuGlnGluLeu	260
QY	994	TTTGATAGATATTCTGGCTATTGGTGTCTCCAAAGCTGAAACAACTCCCAAGTGGTGGT	1053
DB	261	PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly	280
QY	1054	AAAATTCCTTGAATTTCTATATGAAGAAATGATGAATCTGAGGTGGAATATTATTCATGTT	1113
DB	281	LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal	300
QY	1114	ACATCCCCTATGTTGGAACCAAGGAGGCGAGATTCAATTCCTCGTTATCTTAAACACAGTACA	1173
DB	301	ThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThr	320
QY	1174	GCAAATCCTAAAGTCATTTTAAGATGTGAGAAATAATGATGATGCTGGAAGGAAGATC	1233
DB	321	AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle	340
QY	1234	ATAGATCTCATAGATAAGGAACCTAAATTCACCTTTTGAGATTCATATTTGAGGAGTGA	1293
DB	341	IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu	360
QY	1294	TATATTGCCAGAGCTGGATGGACTCCTGAGGGGAAATAATATGCTGTGTCTCATCTACTAGAT	1353
DB	361	TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp	380
QY	1354	CGCTCCCAGACTCGCTACAGATAGTGTGATCTCACCTGGAATTTATTTATCCAGTAGAA	1413
DB	381	ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu	400
QY	1414	GATGATGTTATGGAAGGAGAGACTCATTGAGTCACTGCTGATTTCTGTGACGCCACTA	1473
DB	401	AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu	420
QY	1474	ATTATCTATGAAGAAACAAACAGACATCTGGATAAATATCATGACATCTTTTCATGTTTT	1533
DB	421	IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe	440
QY	1534	CCCCAAGTCACGAAGAGGAAATTTGATTTTATTTTTCCTCTGAATGCAAAACAGGTTTC	1593
DB	441	ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe	460
QY	1594	CGTCAATTATACAAAATTATCATCTATTTTAAAGGAAGCAATATATAACGATCCAGTGGT	1653
DB	461	ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly	480
QY	1654	GGCTGCTGCTCCAGTGAATTTCAAGTGTCTCTATCAAGAGAGAGATGCAATTTACCACT	1713
DB	481	GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer	500
QY	1714	GCTGAATGGGAAGTTCTTGGCCGCATGGATCTAATATCAAGTTGATGAAGTCAGAAAGG	1773
DB	501	GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg	520
QY	1774	CTGGTATATTTTGAAGGCACAAAGACTCCCTCTTTAGAGCATCACCTGTACGTAGTCAGT	1833
DB	521	LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisIleLeuTyrValValSer	540
QY	1834	TACGTAAATCTCGGAGAGGTGACAGGCTGACTGACCGGTGGCTACTCACATCTTTGCTGC	1893
DB	541	TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys	560

```

; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: YUE, Henry
; APPLICANT: ELLIOTT, Vicki
; APPLICANT: GANDHI, Ameena R.
; APPLICANT: LAL, Preeti
; APPLICANT: LAL, YOUNG, Janice
; APPLICANT: TRIBOULEY, Catherine M.
; APPLICANT: DELEGEANE, Angelo M.
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: NGUYEN, Damiel B.
; APPLICANT: LEE, Ernestine A.
; APPLICANT: HAFALIA, April
; APPLICANT: KHAN, Farrah A.
; APPLICANT: CHAWLA, Narinder K.
; APPLICANT: YAO, Monique G.
; APPLICANT: LU, Dyung Aina M.
; APPLICANT: ARVIZU, Chandra S.
; APPLICANT: TANG, Y. Tom
; APPLICANT: WALSH, Roderick T.
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: LU, Yan
; APPLICANT: RAMKUMAR, Jayalaximi
; APPLICANT: XU, Yuming
; APPLICANT: REDDY, Roopa
; APPLICANT: DAS, Depopriya
; APPLICANT: KEARNEY, Liam
; APPLICANT: KALLICK, Deborah A.
; TITLE OF INVENTION: proteases
; FILE REFERENCE: PI-0123 PCT
; CURRENT APPLICATION NUMBER: US/10/311,035
; PRIOR FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: 60/212,336; 60/213,955; 60/215,396; 60/216,821; 60/218,940
; PRIOR FILING DATE: 2000-06-16; 2000-06-22; 2000-06-29; 2000-07-07; 2000-07-14
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PERL Program
; SEQ ID NO 9
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20040023243A1 7160544CD1
; US-10-311-035-9

Alignment Scores:
Pred. No.: 0 Length: 882
Score: 4700.00 Matches: 882
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 84.7% Indels: 0
DB: Gaps: 0

US-10-825-632-2 (1-3120) x US-10-311-035-9 (1-882)
QY 214 ATGCAGCAGCAATGGAAACAGAACAGCTGGGTGTGAGATATTTGAACTGCCGACTGT 273
DB 1 MetAlaAlaMetGluThrGluGluLeuGlyValGluIlePheGluThrAlaAspCys 20
QY 274 GAGCAGGAATATTGAATCAGAGTACGGCTAAATTTGGAGCCCTTTTATTTGTTGAGCGGTAT 333
DB 21 GluGluAsnIleGluSerGlnAspArgProLeuLeuGluProPheTyrValGluArgTyr 40
QY 334 TCCTGGAGTCAGCTTAAAGAGCTGTCTGCCGATACCAGAAAAATATCATCGGTACATGATG 393
DB 41 SerTrpSerGlnLeuLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
QY 394 GCTAAGGCACCACATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATCA 453
DB 61 AlalysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
QY 454 GACAGAACTATTACCTTGGCATGTCTGGTGAGAACAGAGAAATACACTGTTTATTCT 513

```

Db	81	AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer	100
Qy	514	GRAATCCCAAACTATCAATAGACAGCAGTCTTAATGCTCTTGGAAGCCCTTTTG	573
Db	101	GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu	120
Qy	574	GATCTTTTTCAGGCAACACTCGACTATGGAATGTATTCTCGAGAAGAAGAACTATTAAAG	633
Db	121	AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg	140
Qy	634	GAAGAAAACGCAATTCGAACAGTCGGAATTTGCTTTACGATTATCACCAAGGAAGTGA	693
Db	141	GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly	160
Qy	694	ACATTTCTGTTCAAGCCGGTAGTGAATTTATACGTAAAGATGAGAGGCCCAAGGA	753
Db	161	ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly	180
Qy	754	TTTACGCAACAACTTTAAGGCCCAATCTAGTGGAACTAGTTGTCCACATACCGATG	813
Db	181	PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet	200
Qy	814	GATCCAAAATATGCCCCCTGATCCAGACTGGATTGCTTTATACATAGCAACATATT	873
Db	201	AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle	220
Qy	874	TGATATCTAACTCTAACACAGAGAAAGAGACTCACTATGTGCAATAGACTA	933
Db	221	TrpIleSerAsnIleValThrArgGluGluArgArgLeuThrTyrValHisAsnGluLeu	240
Qy	934	GCCAACTGGAAGAGATCCAGATCAGCTGAGAGTCGTACTTTGTTCTCCAAAGAA	993
Db	241	AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu	260
Qy	994	TTTGATAGATATTCTGGCTATTGGTGTCCAAAGCTGAAACAACTCCAGTGGTGT	1053
Db	261	PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly	280
Qy	1054	AAATTTCTAGAAATCTATATGAGAAATGATGAATCTGAGGTGGAAATTTATCATGTT	1113
Db	281	LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHisVal	300
Qy	1114	ACATCCCCTATGTTGGAACAAGGAGGCGAGATTCAATTCGGTTATCTCTAAACAGGTACA	1173
Db	301	ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr	320
Qy	1174	GCAATCCTAACTCACTTTTAAGATGTCAGAAATAATGATTGATGCTGAAGGAAGGATC	1233
Db	321	AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle	340
Qy	1234	ATAGATCTCATAGATAAGAACTAAATTCAACTTTTGAGATTTTATTTGAGGAGTTGAA	1293
Db	341	IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu	360
Qy	1294	TATATTGCCAGAGCTGGAGTCTCCTGAGGAAAATAATGCTGGTCCATCCTACTAGAT	1353
Db	361	TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp	380
Qy	1354	CGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATTTATTTATCCAGTAGAA	1413
Db	381	ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu	400
Qy	1414	GATGATGTTATGAAGGCGAGACTCATTTAGTACGTGCTGATTCTGTGACGCCACTA	1473
Db	401	AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu	420
Qy	1474	ATTATCTATCAAGAAACACAGACATCTCGATAAATATCCATGACATCTTTTCATGTTTTT	1533
Db	421	IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe	440
Qy	1534	CCCCAAAGTCACGAAGGAAATGAGTTTATTTTGTCTCTGAATGCAAAACAGGTTTC	1593
Db	441	ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe	460
Qy	1594	CGTCATTTATACAAAATTACATCTATTTTAAAGGAAAGCAAAATATAAACGATCCAGTGT	1653
Db	461	ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly	480
Qy	1654	GGGCTGCTGCTCCAAGTGATTTCAAGTGTCTCTATCAAGAGGAGATGACAAATACCGT	1713
Db	481	GlyLeuProAlaProSerAspPheLysCysProIleLysGluIleAlaIleThrSer	500
Qy	1714	GGTGAATGGGAAGTTCTTGGCCGGCATGGATCTATATATCCAAAGTTGATGAAGTCAGAAG	1773
Db	501	GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg	520
Qy	1774	CTGTATATTTTCAAGGCCCAAGACTCCCCTTTAGAGACATCACTCTGCTAGCTAGTCAGT	1833
Db	521	LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLysLeuTyrValValSer	540
Qy	1834	TACGTAATCCTCGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACATTTCTTGTGTC	1893
Db	541	TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys	560
Qy	1894	ATCAGTCAGCACTGTGACTTCTTTTATAAGTAAAGTATAGTAACCAAGAAATCCACATGT	1953
Db	561	IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys	580
Qy	1954	GTGTCCTTTTACAAGCTATCAAGTCTGAAGATGACCCCACTTGCAGAAAACAAAGAAATTT	2013
Db	581	ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe	600
Qy	2014	TGGGCCACCACTTTTGGATTTCAGCAGGTCTCTCTCTGACTATACCTCCAGAAATTTTC	2073
Db	601	TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe	620
Qy	2074	TCTTTTCAAGTACTACTCGATTTCATTGTATGATGATGCTCTACAAGCCTCATGATCTA	2133
Db	621	SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu	640
Qy	2134	CAGCTCGAAAGAAATATCTACTGTGCTGTTTATATATGTTGGTGGTCTCAGGTGCGAGTTG	2193
Db	641	GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu	660
Qy	2194	GTGAATTAATCGGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCTCAAGCTCTCTAGT	2253
Db	661	ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly	680
Qy	2254	TATGTGTTGTAGTATAGACAAACAGGGAATCTCTGTACCGAGGCTTAAATTTGAAGGC	2313
Db	681	TyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly	700
Qy	2314	GCCTTTAAATATAAATGGGTCAATAGAAATTTGACGATCAGGTGGAAGGACTCCAATAT	2373
Db	701	AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr	720
Qy	2374	CTAGCTTCTCGATATGATTTCAATTGACTTAGATCGTGGGCATCCAGGCTGGTCCCTAT	2433
Db	721	LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyr	740
Qy	2434	GGAGGATACCTCTCCCTGATGGCATTTAATGCAGAGGTACAGATATCTTCAGGCTTGCTATT	2493
Db	741	GlyGlyTyrLeuSerLeuMetalAlaLeuMetGlnArgSerAspIlePheArgValAlaIle	760
Qy	2494	GCTGGGCCCCAGTCACTCTGTGGATCTTTCTATGATACAGGATACACGGAACGTTTATATG	2553
Db	761	AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet	780
Qy	2554	GGTCACCCCTGACCAAGATGAACAGGCTATTACTTAGGATCTGTGCCCATGCAACAGAA	2613
Db	781	GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu	800
Qy	2614	AAGTTCCCTCTGAAACCAATCGTTTACTGCTTACATGCTTCTCGATGAGAAATGTC	2673
Db	801	LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal	820

```
QY 2674 CANTTTCACATACCATATATCTAGCTTTTGTAGTGGCTGGAAGCCATATGAT 2733
Db 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840

QY 2734 TTACAGATCTATCTCGAGGAGACACAGCATTAAGATTCCTGAATCGGGAGAACATAT 2793
Db 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860

QY 2794 GAACCTGATCTTTTGACATACCTTCAAGAAACCTTGGATCACGTAATGCTGCTCAAAA 2853
Db 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAenLeuGlySerArgIleAlaLeuLeuLys 880

QY 2854 GTGATA 2859
Db 881 ValIle 882

RESULT 5
US-10-072-012-622 Application US/10072012
; Publication No. US20040033493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zernusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Shinkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Esha
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Colman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grose, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 622
; LENGTH: 882

; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-072-012-622

Alignment Scores:
Pred. No.: 0 Length: 882
Score: 4700.00 Matches: 882
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 84.7% Indels: 0
DB: 4 Gaps: 0

US-10-825-632-2 (1-3120) x US-10-072-012-622 (1-882)
QY 214 ATGGCAGCAGCAATGGAAACAGAAACAGCTGGGTGTGGAGATATTTGAACTCCGACTGT 273
Db 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20

QY 274 GAGGAGATATTTGAATCACAGGATCGGCTAAATTTGGAGCCTTTTATGTGTGAGCGGTAT 333
Db 21 GluGluAenIleGluSerGlnAspArgProLysLeuLeuProPheTyrValGluArgTyr 40

QY 334 TCCTGGAGTCAGCTTAAAGAGCTGCTGGCGATACCGAGAAATATCATCGGTACATGATG 393
Db 41 SerTyrSerGlnLeuLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60

QY 394 GCTAAGGCCACCATGATTTTCAATTTGTTGTAAGAGAAATGATCCAGATGACCTCAATCA 453
Db 61 AlaLysAlaProHisAspPheMetPheValLysArgAenAspProAspGlyProHisSer 80

QY 454 GACAGAATCTATTACCTTGGCATGCTGTGGTGAAGACAGAGAAATACACTGTTTATTCT 513
Db 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAenArgGluAenThrLeuPheTyrSer 100

QY 514 GAAATTTCCAAACATATCAATAGACAGCAGTCTTAAATGCTCTCTTTGGAAGCTCTTTTG 573
Db 101 GluIleProLysThrIleAenArgAlaAlaValLeuMetLeuSerTyrPysProLeuLeu 120

QY 574 GATCTTTTTCAGGCAACACTGGACTATGGAATGTATTCTCAGAGAAGAACAATTTAAGA 633
Db 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140

QY 634 GAAGAAACAGCATTTGGAACAGTTCGGAATTCCTTCTTACGATTTATCACCAGGAAGTGA 693
Db 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160

QY 694 ACATTTCTGTTTCAAGCCGCTAGTGGAAATTTTATCAGCTAAAGATGAGAGGCCACAGGA 753
Db 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180

QY 754 TTTACGCAACAACTTTAAGCCCACTAGTGGAAATAGTTCTTCTCCCAACATACGATG 813
Db 181 PheThrGlnGlnProLeuArgProAenLeuValGluThrSerCysProAenIleArgMet 200

QY 814 GATCCAAAATTTATGCCCCGCTGATCCAGACTGGATTCGTTTATATACATACGACGATTT 873
Db 201 AspProLysLeuCysProAlaAspProAspTyrIleAlaPheIleHisSerAenAspIle 220

QY 874 TGGATATCTAAACATCGTAAACAGAGAAGAGAGACTCACTTATGTGCACATGACCTA 933
Db 221 TrpIleSerAenIleValThrArgGluGluArgArgLeuThrTyrValHisAenGluLeu 240

QY 934 GCCAAACATGGAAGAAGATGCCAGATCAGCTGAGTGCCTACTTTGTTCTCCAAAGAGAA 993
Db 241 AlaAenMetGluGluAenAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu 260

QY 994 TTTGATAGATATTTCTGGCTATTTGGTGTGTCCTCAAAAGCTGAAACAACTCCAGTGGTGGT 1053
Db 261 PheAspArgTyrSerGlyTyrTrpTyrCysProLysAlaGluThrThrProSerGlyGly 280

QY 1054 AAAATTTCTAGAATTTCTATATGAAGAAATGATGAATCTGAGGTGGAATTTATTCATGTT 1113
Db 281 LysIleLeuArgIleLeuTyrGluGluAenAspGluSerGluValGluIleIleHisVal 300
```

QY	1114	ACATCCCTATGTTGAACACAGGAGGCGAGATTCAATCCGTTATCTCTAAACAGGTACA	1173
Db	301	ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr	320
QY	1174	GCAATCCTAAAGTCACTTTTAAGATGTCAGAAATAATGATTGATGCTGGAAGGAGGATC	1233
Db	321	AlaAsnProLysValThrPheLysMetSerGluLeuMetIleAspAlaGluGlyArgIle	340
QY	1234	ATAGATGTCATAGATAAGGAACATAATCAACCTTTTGAGATTCTATTGAAAGGAGTTGAA	1293
Db	341	IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu	360
QY	1294	TATATCCAGAGCTGGATGAGTCTCTGAGGGAATAATGCTTGCTGCATCTCTACTAGAT	1353
Db	361	TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp	380
QY	1354	CGTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAAATTTATTTATCCAGTAGAA	1413
Db	381	ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu	400
QY	1414	GATCATGTTTGAAGGACAGACTCATTGAGTCAGTCCCTGATTTCTGTGACGCCACTA	1473
Db	401	AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu	420
QY	1474	ATTATCTATGAAGAAACACAGACATCTGATAAATATCATGACATCTTTTCATGTTTTT	1533
Db	421	IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe	440
QY	1534	CCCCAAAGTCACGAAGGAAATTGAGTTTATTTTGGCTCTGAATGCAAAACAGGTTTC	1593
Db	441	ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe	460
QY	1594	CGTCATTTATACAAATTAATCATCTATTTTAAAGAAAGCAATATAACGATCCAGTGT	1653
Db	461	ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSergly	480
QY	1654	GGGTGCTGCTCCAGTGATTTCAAGTGTCTTCAAGAGAGGATAGCAATACCAAGT	1713
Db	481	GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer	500
QY	1714	GGTCAATGGGAAGTCTTGCGCGCATGGATCTAATCAAGTTGATGAAGTCAGAAGG	1773
Db	501	GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg	520
QY	1774	CTGTTATATTTTGAAGCACCAAGACTCCCTTTTAGAGCATCACCTGTACGTAGTCAGT	1833
Db	521	LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisIleLeuTyrValValSer	540
QY	1834	TACGTAAATCCTGGAGAGGTGACAAGGCTGACTGACCGTGGCTACTCACATTTCTGCTGC	1893
Db	541	TyrValAsnProGlyGluValThrArgLeuThrAspArgLysTyrSerHisSerCysCys	560
QY	1894	ATCAGTCAGACTGTGACTTCTTTAATGATAGTATAGTAAACAGAAATCCACTGT	1953
Db	561	IleSerGlnHisCysAspPheIleSerLysTyrSerAsnGlnLysAsnProHisCys	580
QY	1954	GTGTCCTTTTCAAGCTATCAGTCTCTGAAGATGACCCCACTTGCAGAAACAGGAATTT	2013
Db	581	ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe	600
QY	2014	TGGGCCACATTTTGGATTACAGAGTCTCTCTCTGACTATACTCTCCAGAAATTTTC	2073
Db	601	TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspLysThrProGluIlePhe	620
QY	2074	TCTTTTGAAGTACTACTGGAATTTACATTTGATGGATGCTCTACAGGCTCATGATCTA	2133
Db	621	SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu	640
QY	2134	CAGCCTGGAAAGAAATATCTCTGCTGTTCATATATGTTGGTGTCTCAGGTGCGATTG	2193
Db	641	GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu	660

QY	2194	GTGATAATCGGTTTAAAGAGAGTCAAGTATTTCCGCTTGAATAACCTAGCCTCTCTAGGT	2253
Db	661	ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly	680
QY	2254	TATGTGGTTTAGTATAGACAAACAGGGGATCCTGTCCACGAGGCTTAAATTTGAAGGC	2313
Db	681	TyrValValValValIleAspAsnArgLysCysHisArgGlyLeuLysPheGluGly	700
QY	2314	GCCTTTAATAATAAATGGGTCAATAGAAATTCAGCATCAGGTGGAGGACTCCCAATAT	2373
Db	701	AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr	720
QY	2374	CTAGCTCTCCATATGATTTTCATTGACTTAGATCGTGTGGGCATCCACGGCTGGTCTAT	2433
Db	721	LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyr	740
QY	2434	CGAGGATACCTCTCCCTGATGGCATTAAATGCAGAGGTTCAGATATCTTCAGGGTTCATT	2493
Db	741	GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle	760
QY	2494	GCTGGGGCCCCAGTCACCTCTGTGGATCTTCTATGATACAGATACACGGAACGTTATATG	2553
Db	761	AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet	780
QY	2554	GGTCACCTGACCAGAAATGAACAGGGCTATTACTTAGATCTGTGGCCATGCAAGCAGAA	2613
Db	781	GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu	800
QY	2614	NAGTTTCCCTCTGAACCAAAATCGTTTACTGCTCTTACATGTTTCTCTGGATGAGAAATGC	2673
Db	801	LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal	820
QY	2674	CATTTTGACATACACAGTATATTTACTGAGTTTCTTTAGTGGGCTGGAAGCCCATATGAT	2733
Db	821	HisPheAlaHisThrSerIleLeuLeuSerPheLeuValAlaGlyLysProTyrAsp	840
QY	2734	TTACAGATCTATCTCAGAGAGACACAGCATAGAGTTCTCTGAATCGGGAGAACATTAT	2793
Db	841	LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr	860
QY	2794	GAACTGCATCTTTGCACTACCTTCAAGAAAACCTTGGATCAGCTATTGCTGCTCTAAA	2853
Db	861	GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLys	880
QY	2854	GTGATA 2859	
Db	881	ValIle 882	
RESULT 6			
US-10-415-122-6			
; Sequence 6, Application US/10415122			
; Publication No. US20040053369A1			
; GENERAL INFORMATION:			
; APPLICANT: THE UNIVERSITY OF SYDNEY			
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES			
; FILE REFERENCE: FP15217			
; CURRENT APPLICATION NUMBER: US/10/415,122			
; CURRENT FILING DATE: 2003-08-07			
; NUMBER OF SEQ ID NOS: 8			
; SOFTWARE: PatentIn version 3.1			
; SEQ ID NO 6			
; LENGTH: 882			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
US-10-415-122-6			
Alignment Scores: ---			
Pred. No.: 0			
Score: 4700.00			
Length: 882			
Matches: 882			
Percent Similarity: 100.0%			
Conservative: 0			
Best Local Similarity: 100.0%			
Mismatches: 0			
Query Match: 84.7%			
Indels: 0			
Gaps: 4			
DB:			

US-10-825-632-2 (1-3120) x US-10-415-122-6 (1-882)

QY 214 ATGCGACGACGAAATGGAAACAGACACGCTGGGTGGATGATGATTTGAACTGCGGACTGT 273
DB 1 MetAlaAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
QY 274 GAGGAGAAATTTGAAATCAGAGATCGCCCTAAATTCGAGCCCTTTTATGTTGAGCGGTAT 333
DB 21 GluGluAsnIleGluSerGlnAspA-gProLysLeuGluProPheTy-ValGluArgTy- 40
QY 334 TCCTGGAGTCAGCTTAAAAAGCTGCTTGGCGATACAGAAATATATCATCGCTCATGATG 393
DB 41 SerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTy-HisGlyTy-ValMetMet 60
QY 394 GCTAAGGACCCACATGATTTTCATGTTTGTGAAGGAATGATCCAGATGGACCTCATTTCA 453
DB 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
QY 454 GACAGAATCTATTACCTTGGCATGTCTGGTGGAAACAGAGAAATATACACTGTTTATTCT 513
DB 81 AspArgIleTyTyLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTy-Ser 100
QY 514 GAAATTCCTCAAACTATCAATAGACGACGATCTTAATGCTCTCTTGGAGGCTCTTTTG 573
DB 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSer-TipLysProLeuLeu 120
QY 574 GATCTTTTCAGGCAACACTGGACTATGGAATGTATTCTCGAGAAGAAGAACTATTAGA 633
DB 121 AspLeuPheGlnAlaThrLeuAspTyGlyMetTy-SerArgGluGluGluLeuLeuArg 140
QY 634 GAAAGAAACGCATGGAAACAGCTCGAATTCCTTACGATTATCCAGAGAAAGTGA 693
DB 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTy-AspTy-HisGlnLysSerGly 160
QY 694 ACATTTCTGTTCAAGCCGCTAGTGGAAATTTATACGTAAAGATGGAGGGCCACAAGGA 753
DB 161 ThrPheLeuPheGlnAlaGlySerGlyIleTy-HisValLysAspGlyGlyProGlnGly 180
QY 754 TTTACGCAACAACTTTAAGCCCAATCTAGTGGAAACTAGTTGTCCTCCCAACATACGGATG 813
DB 181 PheThrGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
QY 814 GATCAAAATTTATGCCCCGTGATCCAGACTGGATTGCTTTTATACATAGCAACGATTT 873
DB 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
QY 874 TGGATATCTAACATCGTAACAGAGAAAGAGAGACTCATTTATGTGCACAATGAGCTA 933
DB 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTy-ValHisAsnGluLeu 240
QY 934 GCCAACATGGAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAGAAGAA 993
DB 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu 260
QY 994 TTTGATAGATATTCTGGCTATTGCTGGTGTCCAAAGCTGAAACAACTCCAGTGGTGGT 1053
DB 261 PheAspArgTySerGlyTyTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280
QY 1054 AAAATTTCTAGAATTTATATGAAGAAATGATGATCTGAGCTGGAAATTTATCATGTT 1113
DB 281 LysIleLeuArgIleLeuTyGluGluAsnAspGluLeuValGluIleHisVal 300
QY 1114 ACATCCCTTATGTTGGAACAAAGAGGCGAGATTTCATTCGTTATCTCTAAACAGGTACA 1173
DB 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTy-ProLysThrGlyThr 320
QY 1174 GCAAACTCTAAAGTCACTTTTAAAGATGTCAGAAATATGATTCGTCGAAAGGAGGATC 1233
DB 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
QY 1234 ATAGATGTCATAGTAAGGAACCTAAATTCACCTTTTGGATTTCTATTGTAAGAGGAGTTGAA 1293

DB 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
QY 1294 TATATTTCCAGAGCTGGATGGACTCCTGAGGAAATATCTTGGTCCATCTCTACTAGAT 1353
DB 361 TyrIleAlaAlaGalaGlyTrpThrProGluGlyLysTy-AlaTrpSerIleLeuLeuAsp 380
QY 1354 CCGTCCAGACTCGCTCAGACTAGTGTGATCTCACTGAAATTTATTTATCCAGTAGAA 1413
DB 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
QY 1414 GATGATGTTATGGAAGCAGAGACTCATTTGAGTCACTGCTGCTGATTTCTGTGACGCCACTA 1473
DB 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
QY 1474 ATTATCTATGAAGAAACACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTTT 1533
DB 421 IleIleTyGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
QY 1534 CCCCAGAGTCAGAGAGGAAATTCAGTTTATTTTTCCTCTGAAATCAAAACAGGTTTC 1593
DB 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
QY 1594 CGTCATTTATACAAAATACATCTATTTTAAAGGAAAGCAAAATATAACGATCCAGTGGT 1653
DB 461 ArgHisLeuTy-LysIleThrSerIleLeuLysGluSerLysTy-LysArgSerSerGly 480
QY 1654 GGGCTGCTGCTCCCAAGTATTTCAAGTGTCTTCAAGAGGAGATAGCAATTTACCAGT 1713
DB 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 500
QY 1714 GGTGAATGGAGTTCCTTGGCGCATGGATCTAATATCCAACTGATGAGTCAGCAAGG 1773
DB 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArg 520
QY 1774 CTGGTATATTTTGAAGGCAACCAAGACTCCCTTTAGAGCATCCCTGTACGTAGTACGT 1833
DB 521 LeuValTyPheGluGlyThrLysAspSerProLeuGluHisLeuTy-ValValSer 540
QY 1834 TACGTAATCTCTGGAGAGGTGACAGGCTGACTGACCGTGCTACTCACATTTCTTGCTGC 1893
DB 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTySerHisSerCysCys 560
QY 1894 ATCAGTCAGCACTGTGACTTCTTTATAGTAAGTATATAGTAAACCAAGAAATCCACTGT 1953
DB 561 IleSerGlnHisCysAspPhePheIleSerLysTySerAsnGlnLysAsnProHisCys 580
QY 1954 GTGTCCCTTTACAAGCTATCAAGTCCTGAAGATGACCAACTTTGCCAAAACAAAGGAATTT 2013
DB 581 ValSerLeuTy-LysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
QY 2014 TGGGCCACCATTTTGGATTTCAGCAGGTCTCTCTGACTATATCTCTCAGAAAATTTTC 2073
DB 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyThrProProGluIlePhe 620
QY 2074 TCTTTTGAAGTACTACTGGATTTACATTTGATGGATGCTCTCAAGCCTCATGATCTA 2133
DB 621 SerPheGluSerThrThrGlyPheThrLeuTyGlyMetLeuTyLysProHisAspLeu 640
QY 2134 CAGCTCGGAAAGAAATATCTCTGCTGCTGTTCATATATGTTGCTCCTCAGGTGCGAGTTG 2193
DB 641 GlnProGlyLysLysTy-ProThrValLeuPheIleTyGlyGlyProGlnValGlnLeu 660
QY 2194 GTGAATATCGGTTTAAAGGACTCAAGTATTTCCGCTTGAATACCTAGCTCTCTAGGT 2253
DB 661 ValAsnAsnArgPheLysGlyValLysTy-PheArgLeuAsnThrLeuAlaSerLeuGly 680
QY 2254 TATGTTGTTAGTATAGACAACAGGGGATTCCTGTCCCGAGGGCTTAAATTTGAAGGC 2313
DB 681 TyrValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 700
QY 2314 GCCTTTAAATATAAATCGGTCAATAGAAATTCACGATCAGGTGGAGGACTCCCAATAT 2373
DB 701 AlaPheLysTy-LysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTy- 720

QY 2374 CTAGCTTCTCGATATGATTTTCATTGACTTAGATCGTGTGGGCATCCACGCTGCTCCTAT 2433
Db 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyIlePheTyr 740
QY 2434 GGAGGATACCTCTCCCTCATGGCAATTAATGACAGAGTTCAGATATCTTCAGGGTTGCTATT 2493
Db 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
QY 2494 GCTGGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGGAAGCTTATATG 2553
Db 761 AlaGlyAlaProValThrLeuTyrIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780
QY 2554 GGTCACTCGACCAAGTGAACAGGGCTATTAATCTAGGATCTGTGGCCATGCAAGCAGAA 2613
Db 781 GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
QY 2614 AGTTCCCTCTGAACCAAGCTTTACTGCTCTTACATGTTTCTCGATGAGATGTC 2673
Db 801 LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
QY 2674 CATTTTGACATPACAGATATATCTAGTGTCTTCTAGTGGGTGGAAGCCATATGAT 2733
Db 821 HisPheAlaHisThrSerLeuLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
QY 2734 TTACAGATCTATCTCAGAGAGACACAGACATAAGAGTTCCTGAATCGGAGAACATTAT 2793
Db 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
QY 2794 GAAGTGCATCTTTTGCATCTTCAAGAAACCTTGGATCACGTAATGCTGCTTAAAA 2853
Db 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeuLys 880
QY 2854 GTGATA 2859
Db 881 ValIle 882

RESULT 7
US-10-825-632-1
; Sequence 1, Application US/10825632
; Publication No. US20040191826A1
; GENERAL INFORMATION:
; APPLICANT: ABBOTT, Catherine Anne
; TITLE OF INVENTION: DIPEPTIDYL PEPTIDASES
; FILE REFERENCE: FCSB-100-Div. 1
; CURRENT APPLICATION NUMBER: US/10/825,632
; CURRENT FILING DATE: 2004-04-15
; PRIOR APPLICATION NUMBER: US 10/070,464
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: PCT/AU00/01085
; PRIOR FILING DATE: 2000-09-11
; PRIOR APPLICATION NUMBER: AU PQ5709
; PRIOR FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: AU PQ2762
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-825-632-1
Alignment Scores:
Pred. No.: 0 Length: 882
Score: 4700.00 Matches: 882
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 84.7% Indels: 0
Gaps: 0
DB: 4

US-10-825-632-2 (1-3120) x US-10-825-632-1 (1-882)

QY 214 ATGGCAGCAGCAATGGAAACAGAACAGCTGGGTGTGAGATATTTGAAACTGCGGACTGT 273
Db 1 MetAlaAlaIleAspLeuThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
QY 274 GAGGAGAATATTGAATTCACAGGATCGGCTAAATTTGGAGCCTTTTATTTGAGCGGTAT 333
Db 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
QY 334 TCTGTGAGTCAAGTTAAAGCTGCTGCCGATACCAGAAAATATCATGGTACATGATG 393
Db 41 SerTrpSerGlnLeuLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
QY 394 GCTAAGCACCACATGATTTTCATGTTTGTGAGAGGATGATCAGATGGACCTCATTC 453
Db 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
QY 454 GACAGAATCTATTACCTTCCCATGTCTGTGAGAAACAGAGAAAATACACTGTTTTATTCT 513
Db 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
QY 514 GAAATTTCCCAAACTATCAATAGACAGCAGTCTTAATGCTCTCTTGGAAAGCCTCTTTTG 573
Db 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
QY 574 GATCTTTTTCAGCAACACTGGACTATGGAATGTATCTCAGAGAAAGAACTATTAA 633
Db 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140
QY 634 GAAAGAAAACGCAATGGAAACAGTCGGAATGCTTCTTACGATTTATCACCAGGAAGTGA 693
Db 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
QY 694 ACATTTCTGTTTCAAGCCGTTAGTGAATTTATCAGTAAAGATGGAGGCCACCAAGGA 753
Db 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
QY 754 TTTTACGCAACACCTTTTAAGGCCCAATCTAGTGGAAACTAGTTGTCCCAACATACGGATG 813
Db 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
QY 814 GATCCAAAATTTATGCCCGCTGATCCAGACTGGATTGCTTTTATACATPAGCAAGATTT 873
Db 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
QY 874 TGGATATCTAACATCGTAACCCAGAGAAAGAGAGACTCATTATGTCCACAATGAGCTA 933
Db 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu 240
QY 934 GCCAACATGGAAGAAGATGCCAGATCAGCTGGAGTCGCTACTTTGTTCTCCAGAGAA 993
Db 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu 260
QY 994 TTTGATAGATATTTCTGGCTATTGTTGGTGTCCAAAAGCTGAAACAACTCCAGTGGTGGT 1053
Db 261 PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280
QY 1054 AAAATTTTAGAATCTATATGAAGAAAATGATGAATCTGAGGTGGAATTTTATCATGTT 1113
Db 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300
QY 1114 ACATCCCTATGTTGGAAAACAGGCGCAGATTCATTCCGTTATCTTAAACACAGGTACA 1173
Db 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320
QY 1174 GCAATCCTAAAGTCACTTTTAAAGATGTCAGAAAATAATCATGTATGCTGAAGAGGATC 1233
Db 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
QY 1234 ATAGATGTCATAGATAAGGAACCTAATTCACCTTTTGTAGATTCATTATTTGAAGAGTGA 1293
Db 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360

QY 1294 TATATCCAGAGCTGGATGGACTCCTGAGGAGAAATATGCTTGCTCATCTACTAGAT 1353
Db 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
QY 1354 CGCTCCAGACTCGCCTACAGATAGTGTGATCTCACCTGAATATTATCCAGTAGAA 1413
Db 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
QY 1414 GATCATCTTATGAAAGGAGAGACTCATTGAGTCAGTCCCTGATTCCTGTCAGCCACTA 1473
Db 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
QY 1474 ATTATCTATGAGAAACACAGACATCTGGATAATATCCATGACATCTTCATGTTTTT 1533
Db 421 IleIleTyrGluGlnThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
QY 1534 CCCCAAGTCA CGAAGAGAAATGAGTTTATTTTTCCTCTGATTCGAAACACAGTTTC 1593
Db 441 ProGlnSerHisGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
QY 1594 CGTCATTTATACAAATATACATCTATTTTAAAGGAAAGCAATATATAACGATCCAGTGGT 1653
Db 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
QY 1654 GGGCTGCCTGCCAAGTGATTTCAAGTGTCTTATCAAGAGAGAGATAGCAATATCCAGT 1713
Db 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluIleAlaIleThrSer 500
QY 1714 GGTGAATGGGAAGTTCTTGGCCGCATGGATCTAATATCAAGTTTGATGAAGTCAGAAG 1773
Db 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
QY 1774 CTGGTATATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACCTGTACGTAGTCAGT 1833
Db 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSer 540
QY 1834 TACGTAAATCTCGAGAGGTGACAAAGCTCAGTCGCGTGTACTCACATTTCTTGCTGC 1893
Db 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
QY 1894 ATCAGTCAGACTGTGACTCTTTTATAGTAAGTATAGTAACTAACCCAGAGAAATCCACACTGT 1953
Db 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
QY 1954 GTGTCCTTTACAGCTATCAAGTCCCTGAGATGACCCACTGTCGAAACAAAGCAATTT 2013
Db 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
QY 2014 TGGGCCACATTTTGGATTACAGAGGTCTCTCTCTGACTATATCTCTCAGAAATTTTC 2073
Db 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 620
QY 2074 TCTTTTGAAGTACTACTGGATTTTACATTTGATGGGATGCTCTACAGGCTCATGATCTA 2133
Db 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
QY 2134 CAGCCTGGAAGAAATATCCTACTGTCTGTCTTATATATGCTGTCTCAGGTCAGTTG 2193
Db 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 660
QY 2194 GTGAATAATCCGTTTAAAGGAGTCAAGTATTTCCGCTTGAATACCCCTAGCTCTCTAGT 2253
Db 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680
QY 2254 TATGTGGTTGTAGTATAGACAAACAGGGGATCCTGTCAACGAGGCTTAAATTTGAAGGC 2313
Db 681 TyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly 700
QY 2314 GCCTTTAAATATATATGGTCAAAATAGAAATTCAGCATCAGGTGGAGGACTCCCAATAT 2373
Db 701 AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr 720
QY 2374 CTAGCTTCGATATGATTTTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 2433

Db 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyLysIleGlyTrpSerTyr 740
QY 2434 CGAGGATACCTCTCCCTGATGGCATTAATGCAGAGCTCAGATATCTTCAGGGTTCGTATT 2493
Db 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
QY 2494 GCTGGGCCCCAGTCACTCTGTGATCTTCTATGATACAGGATACACGGAAAGCTTATATG 2553
Db 761 AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780
QY 2554 GGTCAACCTCGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCAGAGAA 2613
Db 781 GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
QY 2614 AAGTTCCCTCTCGAACCAAAATCGTTTACTCTCTTACATCGTTTCTCGATGAGAATGTC 2673
Db 801 LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
QY 2674 CATTTTGCACATACACAGTATATTAATCTAGTGTAGTGAGGCTGGAAAGCCATATGAT 2733
Db 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
QY 2734 TTACAGATCTATCTCAGGAGAGACAGCATTAAGAGTTCTCTGAATCGGAGAACATTA 2793
Db 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
QY 2794 GAAGTGCATCTTTGACACTACCTCAAGAAACCTTGGATCAGTATTCGTCTCTAAAA 2853
Db 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeuLys 880
QY 2854 GTGATA 2859
Db 881 ValIle 882

RESULT 8
US-10-982-512-1
; Sequence 1, Application US/10982512
; Publication No. US2005005081A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/10/982,512
; CURRENT FILING DATE: 2004-11-05
; PRIOR APPLICATION NUMBER: US/09/976,674
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-982-512-1

Alignment Scores:
Pred. No.: 0 Length: 882
Score: 4700.00 Matches: 882
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 84.7% Indels: 0
DB: 5 Gaps: 0

US-10-825-632-2 (1-3120) x US-10-982-512-1 (1-882)

QY 214 ATGGCAGCAGCAATGGAAACAGAACGCTGGGTGTGTGATATTTGAACTCGGACTGT 273
Db 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheLeuThrAlaAspCys 20

QY	274	GAGGAGATATTGAATTCACAGGATCGGCTAAATTTGGAGCCTTTTATGTTGAGCGGTAT	333
DB	21	GlulGluasnilegluSerGlnAspArgProlysleuGluProPheTyrValGluuArgTyr	40
QY	334	TCCTGGAGTCAGCTTAAAGGCTGCTCCGATACACAGAAATATCATGCTCATCATGATG	393
DB	41	SerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet	60
QY	394	GCTAAGGCAACCATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGAGCTCATTTCA	453
DB	61	AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer	80
QY	454	GACAGAACTTATACCTTCCCATGCTCGTGGTGAAGACAGAGAAATACACATGTTTATTCT	513
DB	81	AspArgIleTyrTyrLeuAlaMetSerGlyLysAsnArgGluAsnThrLeuPheTyrSer	100
QY	514	GAAATCCCAAACTATCAATAGACAGCAGCTCTTAATGCTCTCTTGGAGGCTCTTTTG	573
DB	101	GlulIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu	120
QY	574	GATCTTTTTCAGCAACACTGGACTATGGAATGTATTCTCGAGAGAAGAACTATTAAAG	633
DB	121	AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg	140
QY	634	GAAGAAACCGCATTTGGAACAGTCGGAATTCCTTACGATTATCACCAAGGAAGTGA	693
DB	141	GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly	160
QY	694	ACATTTCTGTTTCAAGCCGCTAGTGGAAATTTATCACGTAAAGATGAGGGCCACAAAG	753
DB	161	ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly	180
QY	754	TTTACCCAAACACCTTTAAGCCCAATCTAGTGGAAACTAGTTGCTCCCAACATACCGATG	813
DB	181	PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet	200
QY	814	GATCCAAATTTATGCCGCTGATCCAGACTGGATGCTTTTATACATAGCAACGATATT	873
DB	201	AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle	220
QY	874	TGGATATCTAAACATCGTAACACAGAGAAGAGAGACTCACTATGTGCAACAATGAGCTA	933
DB	221	TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu	240
QY	934	GCCAACTCGAAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCCAAGAGAA	993
DB	241	AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu	260
QY	994	TTTGTATAGATATTCTGGCTATTGGTGTGTCACAAAGCTGAAACAACTCCAGTGGTGT	1053
DB	261	PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly	280
QY	1054	AAAAATCTTTAGAAATCTATATGAAGAAATGATGAATCTGAGGTGGAAATTTATCATGTT	1113
DB	281	LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal	300
QY	1114	ACATCCCTATGTTGGAACAAGGAGGCGAGATTCATTCCGTTATCTCTAAACAGAGTACA	1173
DB	301	ThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThr	320
QY	1174	GCAATCTCTAAAGTCACTTTTAAGATGTCAGAAATATGATGCTGAGGAGCAAGGATC	1233
DB	321	AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle	340
QY	1234	ATAGATGTCATAGATAAGCACTAAATTCACCTTTTGAGATTTCTATTGTAAGAGATTGAA	1293
DB	341	IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu	360
QY	1294	TATATTGCCAGAGCTGGATGGAATCTCGAGGGGAAATATGCTTGGTCCATCTCTACTAGAT	1353
DB	361	TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp	380
QY	1354	CGCTCCAGACTCGCCTACAGATAGTGTGTGATCTCACCTGAATATTATTTATCCCAAGTAA	1413
DB	381	ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu	400
QY	1414	GATGATGTTATGAAAGGCGAGACTCATTTGAGTCAGTGCCTGATTCTGTGACGCCACTA	1473
DB	401	AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu	420
QY	1474	ATTATCTATGCAAGAAACAACACACATCTCGGTAATAATATCCATGACATCTTTTCATGTTT	1533
DB	421	IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe	440
QY	1534	CCCCAAGTCACCAAGCAGGAAATTCAGTTTATTTTTCCTCTGTAATGCAAAACAGGTTTC	1593
DB	441	ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe	460
QY	1594	CGTCATTATACAAAAATTACATCTATTTTAAAGGAAAGCAAAATATAAATCCAGTCCAGTGT	1653
DB	461	ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly	480
QY	1654	GGCTCGCTCTCCAAAGTGAATTTCAAGTGTCTATCAAGAGGAGATAGCAATACCAGT	1713
DB	481	GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer	500
QY	1714	GGTGAATGGGAAGTTCTTGGCGGATGATCTAATATATCCAAAGTTGATGAAGTCAGAAG	1773
DB	501	GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg	520
QY	1774	CTGGTATATTTGAAGGCAACCAAGACTCCCTTTAGAGCATCACCTGTAGTAGTCAGT	1833
DB	521	LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSer	540
QY	1834	TAGTAAATCTCGAGAGGTGACAGGCTGACCTGCGTACTCTACATTTCTTGTGTC	1893
DB	541	TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys	560
QY	1894	ATCAGTCAGCACTGTCACTTCTTATAGTAAGTATAGTACCAAGAGAAATCCACACTGT	1953
DB	561	IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys	580
QY	1954	GTGTCCCTTTACAAAGCTATCAAGTCCTGGAAGATGCCCAACTTGCACAAACAAAGAAATTT	2013
DB	581	ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe	600
QY	2014	TGGGCCCACTTTTGGATTCAGCAGGTCTCTTCTCTGACTATCTCTCCAGAAATTTTC	2073
DB	601	TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePhe	620
QY	2074	TCCTTTGAAAGTACTACTGATTTACATTTGATGGATGGCTCTACAGCCTCATGACTA	2133
DB	621	SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu	640
QY	2134	CAGCTCGGAAGAAATATCTACTGTGTTTCATATATATGTTGGTCTCTCAGTGCACCTG	2193
DB	641	GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu	660
QY	2194	GTGAATAATCGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCTAGCTCTCTAGGT	2253
DB	661	ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly	680
QY	2254	TATGTGGTTGTAGTAGACAAACAGGGGATCTCTGTACCGAGGGCTTTAAATTTTGAAGGC	2313
DB	681	TyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGluGly	700
QY	2314	GCCTTTAAATATATAATGGGTCAAAATAGAAATTTGACGATCAGGTGGGAAGGACTCCAATAT	2373
DB	701	AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr	720
QY	2374	CTAGCTTCGATATGATTTTCATTTAGATCTGTGGGATCCAGGCTGGTCTCTAT	2433
DB	721	LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTrpSerTyr	740
QY	2434	GGAGGATACCTCTCCCTGATGGCATTTAATGCAGAGGTGAGATATCTTCTCAGGTTGCTATT	2493

Db 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
Qy 2494 GCTGGGCCCCAGTCTCTGGATCTTCTATGATACAGGATACAGGAACGTTATATG 2553
Db 761 AlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyrMet 780
Qy 2554 GGTCACTCAGCAGGATGACAGGCTATTACTTATGATCTGTGGCCATCAAGCAGAA 2613
Db 781 GlyHisProAspGlnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
Qy 2614 AAGTCCCTCTGACCAATCTTTACTGCTCTACATGCTTCTGATGAGAGATGC 2673
Db 801 LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
Qy 2674 CATTTGCACATACAGTATATTAATCTAGTCTTTTATGAGGGCTGGAAAGCCATATGAT 2733
Db 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
Qy 2734 TTACAGATCTATCTCTAGGAGACACAGCATTAAGATTCCTGAAATCGGGAGACATTAT 2793
Db 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
Qy 2794 GAATCGATCTTTTGCACCTACCTTCAAGAAACCTTTGGATCAGTATGCTGCTCAAAA 2853
Db 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeuLys 880
Qy 2854 GTGATA 2859
Db 881 Vallie 882

RESULT 9
US-10-072-012-621
; Sequence 621, Application US/10072012
; Publication NO. US20040033493A1
; GENERAL INFORMATION:
; APPLICANT: Ichernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Shinkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Esha
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Colman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grosse, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072, 012
; PRIORITY FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31

; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 621
; LENGTH: 883
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-072-012-621
Alignment Scores:
Prod. No.: 0 Length: 883
Score: 4528.50 Matches: 845
Percent Similarity: 97.7% Conservative: 18
Best Local Similarity: 95.7% Mismatches: 19
Query Match: 81.6% Indels: 1
DB: 4 Gaps: 1
US-10-825-632-2 (1-3120) x US-10-072-012-621 (1-883)
Qy 214 ATGGCAGCAGCATGGAACAGACAGCTGGGTGTGAGATATTTGAAACTCGGACTGT 273
Db 1 MetAlaAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaGluCys 20
Qy 274 GAGGAG--AATATTGAATCAACAGATCGCGCTTAATTTGGAGCTTTTATGTTAGCGG 330
Db 21 GluGluGlyAsnGlyGluSerGlnAspArgProLysLeuGluProPheTyrValGluArg 40
Qy 331 TATTCCTGAGTCAGCTTTAAAAAGCTGCTGCCGATACAGAAATATCATGCTACATG 390
Db 41 TyrSerTrpSerGlnLeuLysLeuAlaAspThrArgTyrHisGlyTyrMet 60
Qy 391 ATGGCTAAGCCACCATGATTTTTCATGTTTGTGAAGAGGAATGATCAGATGACCTCAT 450
Db 61 MetAlaLysAlaProHisAspPheMetPheValLysArgThrAspProAspArgProHis 80
Qy 451 TCAGACAGAAATCTATTACCTTGCCTATGTCGTGAGAACACAGAGAAATACACTGTTTAT 510
Db 81 SerAspArgValTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyr 100
Qy 511 TCTGAAATTCACAAACTATCAATAGACAGCAGTCTTAATGCTCTCTTGGAGCCCTCT 570
Db 101 SerGluIleProLysThrIleAsnArgAlaValLeuMetLeuSerTrpLysProLeu 120
Qy 571 TTGGATCTTTTTCAGCAACACTGGACTATGGAATGTATTCTCGAGAGAGAACTATTA 630
Db 121 LeuAspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeu 140
Qy 631 AGAGAAAGAAACGATTTGGACAGTCCGAAATGCTCTTACGATTATCACCAAGGAAGT 690
Db 141 ArgGluArgLysArgIleGlyThrValGlyIleAlaTyrAspTyrHisProGlySer 160
Qy 691 GGAACATTTCTGTTTCAACCGGTAGTGAATTTATACGTAAAGATGAGAGGCCCAAA 750
Db 161 GlyThrPheLeuPheGlnAlaGlySerGlyIleTyrHisIleLeuAspGlyGlyProHis 180
Qy 751 GGATTTACGCAACACTTTTAAAGGCCCAATCTAGTGGAACTAGTGTCCCAACATACGG 810
Db 181 GlyPheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArg 200
Qy 811 ATGGATCCAAAATATTATGCCCGCTGATCCAGACTGGATTGCTTTTATCATAGCAGCAT 870
Db 201 MetAspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAsp 220

[illegible]

Db	581	CysValSerLeuTyrLysLeuSerSerProGluAspSerProValHisLysThrLysGlu	600
QY	2011	TTTTGGGGCCACCATTITGGATTTCAGCAGGTCCTCTCTCTGACTATACCTACCTCTCCAGAAATT	2070
Db	601	PheTrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIle	620
QY	2071	TTCTCTTTTGAAGTACTACTGGATTTCACATTGTATGGGATGCTCTACAAGCCTCATGAT	2130
Db	621	PheSerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisasp	640
QY	2131	CTACAGCCTGGGAAGAAATATCTTACTGTGCTGTTTCATATATGTGTGTCCTCAGGTGAG	2190
Db	641	LeuGlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGln	660
QY	2191	TTGTGGAAATAATCGGTTTAAAGAGTCAGTATTTCCGCTTGAATACCTACCTAGCCTCTCTTA	2250
Db	661	LeuValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeu	680
QY	2251	GGTTATGTGGTTGTAGTGATAGACACAGGGATCTCTGCACGAGGGCTTAAATTTGAA	2310
Db	681	GlyTyrValValValValIleAspAsnArgGlySerCysHisArgGlyLeuLysPheGlu	700
QY	2311	GGCGCCTTTAAATATAAAATGGGTCAAAATAGAAATTTGACGATCAGTGGAGGACTCCAA	2370
Db	701	GlyAlaPheLysTyrLysMetGlyGlnIleGluLeuAspGlnValGluGlyLeuGln	720
QY	2371	TATCTAGCTTCTCGATATGATTTCATTGACTTAGATCGTGGGCATCCACGGCTGGTCC	2430
Db	721	TyrLeuAlaSerGlnTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrPser	740
QY	2431	TATGGAGGATACCTCCCTCGATGCGCAATTAACACAGGTCAGATATCTCAGGTTGTCT	2490
Db	741	TyrGlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAla	760
QY	2491	ATTGCTGGGGCCCCAGTCACCTCTGTGGATCTTCTATGATACAGGATACACGGAACGTTAT	2550
Db	761	IleAlaGlyAlaProValThrLeuTrpIlePheTyrAspThrGlyTyrThrGluArgTyr	780
QY	2551	ATGGGTACCCCTGCACAGAATGAACAGGGCTATTACTTAGGATCTGTGCCCATCGAAGCA	2610
Db	781	MetGlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAla	800
QY	2611	GAAGAATCCCTCTGTAACACCAATTCGTTTACTGCTCTTACATGGTTTCTCGATGAGAAAT	2670
Db	801	GluLysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsn	820
QY	2671	GTCCATTTTGCAATACCAGTATATTACTGAGTTTTTTAGTGAGGGCTGGAAAGCCATAT	2730
Db	821	ValHisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyr	840
QY	2731	GATTACAGATCTATCTCCAGAGAGACACAGCATTAAGAGTTCCTGAATCCGGGAGACAT	2790
Db	841	AspLeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHis	860
QY	2791	TATGAATCGCATCTTTTCGACTACCTTCAAGAAAACCTTGGATCAGGTATTCCTGCTCTTA	2850
Db	861	TyrGluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaAlaLeu	880
QY	2851	AAAGTCATA	2859
Db	881	LysValIle	883

RESULT 10
US-09-976-674-7
; Sequence 7, Application US/09376674
; Patent No. US20020115843A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV

FILE REFERENCE: 70669		CURRENT APPLICATION NUMBER: US/09/976,674	
CURRENT FILING DATE: 2001-10-12		PRIOR APPLICATION NUMBER: US 60/240,117	
PRIOR FILING DATE: 2000-10-12		NUMBER OF SEQ ID NOS: 61	
SOFTWARE: PatentIn version 3.1		SEQ ID NO 7	
LENGTH: 690		TYPE: PRT	
ORGANISM: Homo sapiens		ORGANISM: Homo sapiens	
US-09-976-674-7			
Alignment Scores:		Length: 690	
Pred. No.: 0		Matches: 690	
Score: 3607.50		Conservative: 0	
Percent Similarity: 87.1%		Mismatch: 0	
Best Local Similarity: 87.1%		Indels: 102	
Query Match: 65.0%		Gaps: 1	
DB: 3			
US-10-825-632-2 (1-3120) x US-09-976-674-7 (1-690)			
QY	214	ATGGCAGCAGCAATGGAAACAGACAGCTGGGTGTTGAGATATTTCAGAACTCCGGACTGT	273
DB	1	MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys	20
QY	274	GAGGAGATATTGAATCACAGGATCGGCCCTAAATTTGGAGCCTTTTATGTTGAGCGGTAT	333
DB	21	GlulGluLeuIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr	40
QY	334	TCCTGGAGCTGAGCTTAAAGAGCTGCTTGGCGATACAGAAATATCATGGCTACATGATG	393
DB	41	SerTrpSerGlnLeuLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet	60
QY	394	GCTAAGGACACATGATTTTCATCTTTGTGAAGAGGAATCATCCAGATGGACTCATTTCA	453
DB	61	AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer	80
QY	454	GACAGAATCTATTACCTTGCATCTGCTGGTGAGAACAGAGAAATACACTGTTTATTCT	513
DB	81	AspArgLysTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer	100
QY	514	GAATTTCCAAACATCATATAGACAGCAGCTCTTAATGCTCTTTGGAAGCCTCTTTTG	573
DB	101	GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu	120
QY	574	GATCTTTTCAGGCAACATCGACTATGGAAATGTTATTCGAGAAGAAGACTATTAGA	633
DB	121	AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg	140
QY	634	GAAGAAAACGATTTGGAACAGTCCGAATTCCTTTACGATTATCACCAAGAGAGTGA	693
DB	141	GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly	160
QY	694	ACATTTCTGTTTCAAGCCGGTAGTGGAAATTTATCACGTAAAGATGGAGGCCACAGA	753
DB	161	ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly	180
QY	754	TTTACCCAAACACCTTTAAGGCCCAATCTAGTGGAACTAGTTGCTCCCAACATACGGATG	813
DB	181	PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet	200
QY	814	GATCCAAATTTATGCCCGCTGATCCAGACTGGATTGCTTTTATACATACCAACGATATT	873
DB	201	AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle	220
QY	874	TGGATATCTAACATCTGATACAGAGAGAAAGGAGACTCACTTATGTGCACAATGAGCTA	933
DB	221	TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu	240
QY	934	GCCAACTGGAAGAAGATCCAGATCAGCTGAGTCCGTACCTTTGTTCTCCAGAGAA	993

Db 361 TyrllealalargalaglyTrpThrProgluGlyLysTyAlaTrpSerlleLeuLeuAsp 380
Qy 1354 CGTCCAGACTCGCTACAGATPAGTGTGATCTCACTCGAATATTATTATCCCAAGTAGAA 1413
Db 381 ArgSerGlnThrArgLeuGlnlleValLeuIleSerProGluLeuPheIleProValGlu 400
Qy 1414 GATGATCTTATGAAAGGACAGACTCATTTGAGTCAGTGCCTGTGATCTTGTGACCCACTA 1473
Db 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
Qy 1474 ATTATCTATGAAGAAACAACAGACATCTGATAAATATCCATGACATCTTTCTGTTT 1533
Db 421 IleIleTyrgluGluThrThrAspIleTrrIleAsnIleHisAspIlePheHisValPhe 440
Qy 1534 CCCAAAGTCACGAAGAGGAAATTTGAGTTTATTTTTGCTCTGAAATGCAAAACAGGTTTC 1593
Db 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
Qy 1594 CGTCATTTATACAAAATTACATCTATTTTAAAGGAAGCAAAATATAAAGCATCCAGTGGT 1653
Db 461 ArgHisLeuTyrlsIleThrSerlleLeuLysGluSerLysTyrlsArgSerSerGly 480
Qy 1654 GGGTGGCTGCTCCAGTGATTTCAAGTGTCTCTATCAAGAGAGATAGCAATACCAAGT 1713
Db 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 500
Qy 1714 GGTGAATGGGAAGTCTTGGCCGCGCATGATCTTAATCCAGTTGATGAGTCAGAGG 1773
Db 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
Qy 1774 CTGGTATATTTGAAGGACCAAGACTCCCTTTAGAGCATCACCTGTACGTAGTCAGT 1833
Db 521 LeuValTyrlPheGluGlyThrLysAspSerProLeuGluHisLeuTyrlValSer 540
Qy 1834 TACGTAAATCTGGAGAGGTGACAAGGCTGACTGACCGTGGCTGCTACTCATCTTTGCTGC 1893
Db 541 TyrlValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrlSerHisSerCys 560
Qy 1894 ATCACTGACAGCTGTGACTCTTTTATAAGTAAGTATAGTAACACAGAAGAAATCCACACTGT 1953
Db 561 IleSerGlnHisCysAspPhePheIleSerLysTyrlSerAsnGlnLysAsnProHisCys 580
Qy 1954 GTGTCCCTTTACAAGCTATCAAGTCTCTGAAGATGACCAACTTGCAAAACAAAGGAATTT 2013
Db 581 ValSerLeuTyrlLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
Qy 2014 TGGGCCACCATTTTGGATTTCAGCAGGTCCTCTTCCTGACTATACTCTCCAGAAATTTTC 2073
Db 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrlThrProGluIlePhe 620
Qy 2074 TCTTTTGAAGTACTGATTTACATTTGATGGGATGCTCTACAGGCTCATGATCTA 2133
Db 621 SerPheGluSerThrThrGlyPheThrLeuTyrlGlyMetLeuTyrlLysProHisAspLeu 640
Qy 2134 CAGCCTGGAAAGAAATATCTACTGCTGTGCTTATATATGCTGCTCAGGTGCGAGTTG 2193
Db 641 GlnProGlyLysLysTyrlProThrValLeuPheIleTyrlGly----- 655
Qy 2194 GTGAATAATCGGTTTAAAGGAGTCAGTAATTTCCGCTTGAATACCCCTAGCCTCTAGGT 2253
Db 655 ----- 655
Qy 2254 TATGTGGTTGTAGTAGACACACAGGGGATCTGTCAACGAGGGCTTAAATTTGAAGC 2313
Db 655 ----- 655
Qy 2314 GCCTTTAAATATAAAATGGGTCAATAGAAAATTGACGATCAGGTGGAGGACTCCAAATAT 2373
Db 655 ----- 655
Qy 2374 CTAGCTTCTCGATGATGATTTCAATTCATTTAGATCGTGTGGGCATCCACGGCTGGTCTAT 2433
Db 655 ----- 655

Qy 2434 GGAGGATACCTCTCCCTGATGGCATTAAATGCAGAGGTGAGATATCTTACAGGCTTGCTATT 2493
Db 656 -----Arg-LeuLeuLe 659
Qy 2494 GCTGGGGCCCGACGTCACCTGTGATCTTCTATGATACAGGATACACGGAACGTTATATG 2553
Db 659 uleuGlyProGlnSerLeuGlySerSerMetIleGlnAspThrArgAsnValIleTr 679
Qy 2554 GGTACCCCTGACCAAGATGAACAGGCTATTACT 2587
Db 679 pValThrLeuThrArgMetAsnArgAlaIleThr 690

RESULT 12
US-09-976-674-11
; Sequence 11, Application US/09976674
; Patent No. US20020115843A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/09/976.674
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 661
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-11

Alignment Scores:
Pred. No.: 1,11e-311 Length: 661
Score: 3513.50 Matches: 660
Percent Similarity: 99.7% Conservative: 0
Best Local Similarity: 99.7% Mismatches: 0
Query Match: 63.3% Indels: 2
DB: 3 Gaps: 1

US-10-825-632-2 (1-3120) x US-09-976-674-11 (1-661)

Qy 214 ATGGCAGCAGCAATCGAAACAGACAGCTGGGTGTGAGATATTTGAACTGCGGACTGT 273
Db 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
Qy 274 GAGGAGAATATTGAATCAACAGATCGGCTAAATTTGAGCCCTTTTATTTGTTGACCGTAT 333
Db 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrlValGluArgIyr 40
Qy 334 TCCTGGAGTCAGCTTAAAGGCTGCTTCCGATACCGAAATATCATGCGGTACATGATG 393
Db 41 SerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrlHisGlyTyrlMetMet 60
Qy 394 GCTAAAGCCACCATGATTTTCATGTTGTGAAGAGGAATGATCCAGATGGACCTCATTC 453
Db 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
Qy 454 GACAGATCTATTACCTTCCCATGCTGTGTGAGACAGACAGAAATATACATCTTTTATCT 513
Db 81 AspArgIleTyrlTyrlLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrlSer 100
Qy 514 GAAATCCCAAACTACTAATAGCAGCAGCTTAAATGCTCTCTTGTGAAGCCTCTTTTG 573
Db 101 GluIleProLysThrIleAsnArgAlaValLeuMetLeuSerTrpLysProLeuLeu 120
Qy 574 GATCTTTTTCAGGCACATCGGACTATGGAATGATTTCTCCGAGAGAGAACTTAAAGA 633
Db 121 AspLeuPheGlnAlaThrLeuAspTyrlGlyMetTyrlSerArgGluGluLeuLeuArg 140

QY	634	GAAGAAACGCAATGGAAACAGCTCGGAATTCCTCTTACGATTATACCAAGGAAGTGA	693	QY	1714	GGTGAATGGGAAGTCTTGGCCGGCATGGATCTAATATCCAAAGTTGATGAAGTCAGAAG	1773
DB	141	GUUAGUyAargllelyThrValGlyllealeaserYzAspYrHisGlnGlySerGly	160	DB	501	GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg	520
QY	694	ACATTTCTCTTTCAAGCCGGTAGTGGAAATTTATCACGTAAGAAGATGAGAGGCCACAAGGA	753	QY	1774	CTGGTATATTTGAAGGCAACCAAGACTCCCCCTTTAGAGCATCACCTGTACGTAGTCAGT	1833
DB	161	ThrPheLeuPheGlnAlaGlySerGlylleYrHisVallyAspGlyGlyProGlnGly	180	DB	521	LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisHisLeuTyrValValSer	540
QY	754	TTTACCCAAACAACCTTTAAGGCCCAATCTAGTGGAAACTAGTTGTCGCCAACATCCGATG	813	QY	1834	TAGTAAATCTCGAGAGGTGCAAGGCTGACTGACCGTGGCTACTCACATTCTTGCTGC	1893
DB	181	PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet	200	DB	541	TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys	560
QY	814	GATCCAAAATTTATGCCCGCTGATCCAGACTGGATTGCTTTTATACATACACATATTT	873	QY	1894	ATCAGTCAGCACTGTGACTTCTTTTAACTAAGTATAGTAACAGAGAATCCACACTGT	1953
DB	201	AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle	220	DB	561	IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys	580
QY	874	TGGATATCTAAATCGTAACAGAGAGAGAGAGACTCACTTATGTGCAATAGAGCTA	933	QY	1954	GTGTCCCTTTACAGCTATCAAGTCTGAAGATGACCAACTTGCAAAACAAAGAAATTT	2013
DB	221	TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu	240	DB	581	ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe	600
QY	934	GCCAAATCGAAGAGATCCAGATCAGCTGGAGTCGCTACCTTTGTCTCCAAAGAA	993	QY	2014	TGGGCCACCACTTTTGAATTCAGAGTCTCTCTTCTGACTATATCTCTCCAGAAATTTTC	2073
DB	241	AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu	260	DB	601	TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe	620
QY	994	TTTGTAGATATTTCTGGCTATTGTGGTGTCCAAAAGCTGAAACAACTCCACGTGGTGT	1053	QY	2074	TCCTTTGAAGTACTACTGATTACATTTGATGGATGGATGCTCTACAAGCCTCATGATCTA	2133
DB	261	PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly	280	DB	621	SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu	640
QY	1054	AAATTTCTTAGAATCTCATATGAAGAAATGATGAATCTGAGGTGGAAATTTATTCATGTT	1113	QY	2134	CAGCTCGAAGAAATATCTACTGCTGCTGTTTCATATATGGTGGTCTC---AGTGCAG	2190
DB	281	LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHisVal	300	DB	641	GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGly-LeuLeuArgCysase	660
QY	1114	ACATCCCTATGTGGAAACAAGAGGCGCAGATTCATTCCTGTTATCTCTAAAACAGGTACA	1173	QY	2191	TTGG 2194	
DB	301	ThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThr	320	DB	660	ITrp 661	
QY	1174	GCAATCTTAAGTCACTTTTAAGATGTCAGAAATATGATGATGCTGAGGAGGATC	1233				
DB	321	AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle	340				
QY	1234	ATGATGTCATAGTAAGCACTAATTCACCTTTGAGATTCCTATTGAGAGAGTTGAA	1293				
DB	341	IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu	360				
QY	1294	TATATTGCCAGACTGGATGACTCTGAGGGAATATGCTGCTGCATCTACTAGAT	1353				
DB	361	TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp	380				
QY	1354	CGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATTTATTTATCCAGTAGAA	1413				
DB	381	ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu	400				
QY	1414	GATGATGTTATGGAAGGAGAGACTCAATGAGTCAGTGCCTGATCTGTGAGCCCACTA	1473				
DB	401	AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu	420				
QY	1474	ATTATCTATGAAGAAACAACAGACATCTGGATAAATATCCATGACATCTTTCATGTTTT	1533				
DB	421	IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe	440				
QY	1534	CCCCAAGTCACAGAGAAATGAGTTTATTTTGGCTCTGATGCAAAACAGGTTTC	1593				
DB	441	ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe	460				
QY	1594	CGTCATTTATACAAATACATCTATTTTAAAGGAAGCAATATATAACCATCCAGTGT	1653				
DB	461	ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly	480				
QY	1654	GGGCTGCTCTCCAAGTGAATTTCAAGTGTCTATCAAAAGAGGAGATAGCAATTAACAGT	1713				
DB	481	GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer	500				
<hr/>							
RESULT 13							
US-10-982-512-11							
; Sequence 11, Application US/10982512							
; Publication No. US20050059081A1							
; GENERAL INFORMATION:							
; APPLICANT: Qi, Steve							
; APPLICANT: Akineanya, Karen							
; APPLICANT: Riviere, Pierre							
; APPLICANT: Junien, Jean-Louis							
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV							
; FILE REFERENCE: 70669							
; CURRENT APPLICATION NUMBER: US/10/982,512							
; CURRENT FILING DATE: 2004-11-05							
; PRIOR APPLICATION NUMBER: US/09/976,674							
; PRIOR FILING DATE: 2001-10-12							
; PRIOR APPLICATION NUMBER: US 60/240,117							
; PRIOR FILING DATE: 2000-10-12							
; NUMBER OF SEQ ID NOS: 61							
; SOFTWARE: PatentIn version 3.1							
; SEQ ID NO 11							
; LENGTH: 661							
; TYPE: PRT							
; ORGANISM: Homo sapiens							
US-10-982-512-11							
<hr/>							
Alignment Scores:							
Pred. No.:	1,11e-311	Length:	661				
Score:	3513.50	Matches:	660				
Percent Similarity:	99.7%	Conservative:	0				
Best Local Similarity:	99.7%	Mismatches:	0				
Query Match:	63.3%	Indels:	2				
DB:	5	Gaps:	1				
<hr/>							
US-10-825-632-2 (1-3120) x US-10-982-512-11 (1-661)							
QY	214	ATGCGACGACCAATGGAAACAGAACAGCTGGGTGGTGTGAGATATTTGAACACTCGGACTGT	273				

Db	1	MetAlaAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys	20
Qy	274	GAGGAGAAATATTGAATCACAGGATCGGCTAAATTTGGAGCCCTTTTATGTGTGAGCGGTAT	333
Db	21	GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr	40
Qy	334	TCTCGGAGTCAGCTTAAAGCTGCTTGCCTGATACAGAGAAATATCATGGCTACATGATG	393
Db	41	SerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet	60
Qy	394	GCTAAGGCACACATGATTTTCATGTTTGTGAAGAGAAATGATCCAGATGGACCTCATTTCA	453
Db	61	AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer	80
Qy	454	GACGAATCTTATACCTTGCCTGCTGCTGAGAGAGAAATGATCCAGATGGACCTCATTTCT	513
Db	81	AspArgIleTyrThrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer	100
Qy	514	GAATTTCCAAAACCTATCATAGAGCAGCGCTTAAATGCTCTCTTGGAGCCCTTTTG	573
Db	101	GluIleProLysThrIleAsnArgAlaValLeuMetLeuSerTyrLysProLeuLeu	120
Qy	574	GATCTTTTCAGGCAACACTGGACTATGGAATGCTTCTTACGATATCACCAAGGAAGTGA	633
Db	121	AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg	140
Qy	634	GAAGAAACACCATTTGGAACAGATCGGAATTTGCTTCTTACGATATCACCAAGGAAGTGA	693
Db	141	GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly	160
Qy	694	ACATTTCTGTTTCAAGCCGCTAGTGGAAATTTATCACGTAAAGATGGAGGCGCACAGGA	753
Db	161	ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly	180
Qy	754	TTTACGCAACACCTTTTAAAGCCCAATCTAGTGGAACTAGTTGTTCCCAACATACGGATG	813
Db	181	PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet	200
Qy	814	GATCCAAATATTATGCCCCGCTGATCCAGACTGGATGCTTTTATATACATGCAACGATAT	873
Db	201	AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle	220
Qy	874	TGGATATCTACATCGTAACACAGAGAAGAAGAGACTCACTTATGTCACACATGAGTGA	933
Db	221	TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu	240
Qy	934	GCAACATGGAAGAGATGCCAGATCAGCTGGAGTCGCTACCTTGTCTTCTCCAAGAAGAA	993
Db	241	AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGlu	260
Qy	994	TTTGATAGATATCTGGCTATTGTTGGTTGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1053
Db	261	PheAspArgTyrSerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly	280
Qy	1054	AAAAATCTTAGAATCTTATATGAAGAAATGATGAATCTGAGTGGAAATTTATCATGTT	1113
Db	281	LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal	300
Qy	1114	ACATCCCTTATGTTGGAACACAGGAGGCGAGATTCATTCCTGTTATCTTAAACAGGTACA	1173
Db	301	ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr	320
Qy	1174	GCAATCTTAAAGTCACTTTTAAAGATGTGAGAAATATGATGATGCTGAAGAGGATC	1233
Db	321	AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle	340
Qy	1234	ATAGATGTCATAGTAAGGAACCTAATTCACCTTTTGAGATTCCTATTGAGGAGTGA	1293
Db	341	IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu	360
Qy	1294	TATATTGCCAGAGCTGGAGCTCCTGAGGAGAAATATGCTTGGTCCATCTCTACTAGAT	1353
Db	361	TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaIleTrpSerIleLeuLeuAsp	380
Qy	1354	CGCTCCAGACTCGCTACAGATAGTGTGATCTCACCTGAATATTATTATCCAGTAGAA	1413
Db	381	ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu	400
Qy	1414	CATCATCTTATGAAAGGACAGACTCATTGAGTGCCTGATTCCTGATTCAGCCACTA	1473
Db	401	AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu	420
Qy	1474	ATTATCTATGAAAGAACACAGACATCTGATAAATATCCATGACATCTTTTCATCTTTT	1533
Db	421	IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe	440
Qy	1534	CCCAAGTCAAGAAAGAAATTTGATTTATTTTCTCTCTCTCAATGCAAAACAGTTTC	1593
Db	441	ProGlnSerHisGluGluGluIleGluPheIleAlaSerGluCysLysThrGlyPhe	460
Qy	1594	CGTCATTTTATACAAAATTTACATCTTATTTTAAAGAAAGCAATATTAACATCCAGTGT	1653
Db	461	ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly	480
Qy	1654	GGCTGCTGCTCCCAAGTGAATTTCAAGTGTCTTATCAAGAGAGATAGCAATATTACAGT	1713
Db	481	GlyLeuProAlaProSerAspPheLysCysProIleLysGluIleAlaIleThrSer	500
Qy	1714	GGTGAATCGGAAGTCTTTGGCCGCGCATGATCTAATATCAAGTTGATGAAGTCAGAAAG	1773
Db	501	GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg	520
Qy	1774	CTGCTATATTTTGAAGGCACCAAGACTCCCTTTTAGAGCATCACCTGTACGTAGTCACT	1833
Db	521	LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisIleLeuTyrValValSer	540
Qy	1834	TACGTAATCTCGAGAGGTCACAAAGCTGACCTGACCTGACCTGGCTACTCACATTTCTGCTG	1893
Db	541	TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys	560
Qy	1894	ATCAGTCAGACTGTGACTCTCTTATTAAGTAAGTATAGTAAACCAGAGAATCCACACTGT	1953
Db	561	IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys	580
Qy	1954	GTGTCCTTTTCAAGCTATCAAGTCTCTGAGAGTACCCTGCAAACTTGCAGAAAGGAATTT	2013
Db	581	ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe	600
Qy	2014	TGGGCCACATTTTGGATTCAGAGTCTCTCTCTGACTATCTCTCCAGAAATTTTC	2073
Db	601	TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe	620
Qy	2074	TCTTTTGAAGTACTACTGGATTTTACATTTGTATGGATGCTCTTACAAGCCTCATGATCTA	2133
Db	621	SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu	640
Qy	2134	CAGCTCGGAAAGAAATATCCTACTGTGCTGTTTCATATATGTTGCTCTCTCTCTCTCT	2190
Db	641	GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyLeuLeuArgCysSe	660
Qy	2191	TTGG 2194	
Db	660	ITP 661	
RESULT 14			
US-09-976-674-19			
; Sequence 19, Application US/09976674			
; Patent No. US20020115843A1			
; GENERAL INFORMATION:			
; APPLICANT: Qi, Steve			
; APPLICANT: Akinsanya, Karen			
; APPLICANT: Riviere, Pierre			
; APPLICANT: Junien, Jean-Louis			
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV			
; FILE REFERENCE: 70669			
; CURRENT APPLICATION NUMBER: US/09/976,674			

; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
; LENGTH: 658
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-674-19

Alignment Scores: 8.17e-311 Length: 658
Pred. No.: 3504.00 Matches: 655
Score: 3504.00
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 63.1% Indels: 0
DB: 3 Gaps: 0

US-10-825-632-2 (1-3120) x US-09-976-674-19 (1-658)

QY 214 ATGCGACGACGAGTGAACAGACAGCTGGGTGTGAGATATTTGAACTGCGGACTGT 273
DB 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
QY 274 GAGGAGAATATTGAATACAGGATCGGCTAAATTTGGAGCCCTTTTATGTTGAGCGGTAT 333
DB 21 GluGluAsnIleGluSerGlnAspArgProlysLeuGluProPheTyrValGluA-GTyr 40
QY 334 TCCTGGAGTCAGCTAAAAAGCTGCTGCCGATACCGAATAATATCATGCTCATCATG 393
DB 41 SerTrpSerGlnLeuIleGlyLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
QY 394 GCTAAGCCACCATGATTTTCATGTTGTGAAGAGAAATGATCCAGATGAGACCTCATTC 453
DB 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
QY 454 GACAGAACTTATTCCTGCCATGCTGCTGTGAGAACAGAGAAATATACACTGTTTATTCT 513
DB 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
QY 514 GAAATTCCTCAAACTATCAATAGACGACGCTTAATGCTCTCTGGAAGCCTCTTTTG 573
DB 101 GluIleProLysThrIleAsnArgAlaValLeuMetLeuSerTrpLysProLeuLeu 120
QY 574 GATCTTTTTCAGGCAACACTGGACTATGAAATGATTTCTCGAGAGAGAACTATTAAAGA 633
DB 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluGluLeuLeuArg 140
QY 634 GAAAGAAAACCGCATTCGAAACAGTCGGAATTTGCTTTACGATTATCACCAAGGAAGTGA 693
DB 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
QY 694 ACATTTCTGTTCAAGCCGGTAGTGAATTTATACGTAAAGATGAGAGGCGCCACAGGA 753
DB 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
QY 754 TTACGCAACAACTTTAAGGCCAATCTAGTGGAACTAGTTGTGCCACATACCGATG 813
DB 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
QY 814 GATCCAAAATATGCCCGCTGATCCAGCTGGATGCTTTTATACATAGCAAGATATT 873
DB 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
QY 874 TGGATATCTAATCTTAACCGAGAGAGAAAGAGACTCACTTATGTGCAATAGAGCTA 933
DB 221 TrpIleSerAsnIleValThrArgGluGluArgGluThrTyrValIleAsnGluLeu 240
QY 934 GCCAATATGAAGAAGATCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCAAAGAGAA 993
DB 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaIleThrPheValLeuGlnGlu 260

QY 994 TTTGATAGATATTTCTGGCTATTGGTGGTGTCAAAGAGCTGAAACAACTCCACAGTGTGGT 1053
DB 261 PheAspArgTyr-SerGlyTyrTrpTrpCysProLysAlaGluThrThrProSerGlyGly 280
QY 1054 AAAATTTCTAGAAATTTATATGAGAAAATGATGAATCTGAGGTGGAAATTTATTCATGTT 1113
DB 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleHisVal 300
QY 1114 ACATCCCTCTATGTTGAAAACAGGAGGCGAGATTCATTCCGTTATCTTAAACAGAGTACA 1173
DB 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320
QY 1174 GCAAAATCCTAAAGTCATCTTTTAAGATGTGAGAAATAATGATTCATTCGTGAAGAGGATC 1233
DB 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
QY 1234 ATAGATGTCATAGATAAGGAACCTAAATTCACCTTTTGTAGATTTCTATTTTGAAGGAGTTGAA 1293
DB 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
QY 1294 TATATGCCAGACTCGATGAGTGAAGCTCTGAGGAGAAATAATGCTGTGCTCATCTACTAGAT 1353
DB 361 TyrIleAlaArgAlaGlyTrpThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
QY 1354 CGTCCGAGACTCGCTACAGATAGTGTGATCTCACCTGAAATTTATTTATCCAGTAGAA 1413
DB 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
QY 1414 GATGATGTTATGAAAAGGAGAGACTCATGTTGAGTCAGTGCCTGATTTCTGTCAGCCACTA 1473
DB 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
QY 1474 ATTATCTATGAGAAACAAACAGACATCTGGATAAATATCCATGATCATCTTTTCATGTTTT 1533
DB 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
QY 1534 CCCCAGAAAGTCACGAGAGGAAATTTGAGTTTATTTTGGCTCTGAAATGCAAAACAGGTTTC 1593
DB 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
QY 1594 CGTCATTTATACAAAATTACATCTATTTTAAAGGAAAGCAATAATAAAGCATCCAGTGGT 1653
DB 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
QY 1654 GGGCTGCTCTCCAAAGTGATTTCAAGTGTCTATCAAGAGGAGATAGCAATACCACT 1713
DB 481 GlyLeuProAlaProSerAspPheLysCysProLysGluGluIleAlaIleThrSer 500
QY 1714 GGTGAATGGGAAGTTCTTGGCCGGCATGATCTAATATCCAAAGTTGATGAAGTCAGAGG 1773
DB 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
QY 1774 CTGCTATATTTTGAAGSCCAAGAGACTCCCTTTAGAGCATCACCTGTAGCTAGTCACT 1833
DB 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisIleLeuTyrValValSer 540
QY 1834 TACCTAAATCCTCGAGAGGTGACAAGCTGACTGACCGGTGCTACTCACATTTCTTGTGTC 1893
DB 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
QY 1894 ATCAGTCAGCACTGTGATCTTTTATTAAGTAAGTATAGTAACAGAGAATCCACACTGT 1953
DB 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
QY 1954 GTGTCCTTTTACAGCTATCAAGTCTCTGAAGATGACCCCACTTGCAGAAAACAGAAATTT 2013
DB 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
QY 2014 TGGGCGCCACCATTTTGGATTTCAGAGTCTCTCTCTCTGACTATACCTCTCCAGAAATTTTC 2073
DB 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 620
QY 2074 TCTTTTGAAGTACTACTGGAATTTACATTTGATGGATGCTCTTACAAGCTCATCATCTA 2133

Db 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
QY 2134 CAGCTCGAAGAAATCTCTACTGTGCTGTTTCATATATGGTGGT 2178
Db 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGly 655
RESULT 15
US-10-982-512-19
; Sequence 19, Application US/10982512
; Publication No. US20050059081A1
; GENERAL INFORMATION:
; APPLICANT: Qi, Steve
; APPLICANT: Akinsanya, Karen
; APPLICANT: Riviere, Pierre
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: NOVEL SERINE PROTEASE GENES RELATED TO DPPIV
; FILE REFERENCE: 70669
; CURRENT APPLICATION NUMBER: US/10/982,512
; PRIOR FILING DATE: 2004-11-05
; PRIOR APPLICATION NUMBER: US/09/976,674
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/240,117
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
; LENGTH: 658
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-982-512-19

Alignment Scores:
Pred. No.: 8,17e-311 Length: 658
Score: 3504.00 Matches: 655
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 63.1% Indels: 0
DB: 5 Gaps: 0

US-10-825-632-2 (1-3120) x US-10-982-512-19 (1-658)

QY 214 ATGGCAGCAGCATCGAAGAAACAGACAGCTGGTGTTCAGATATTGCAACTGGGACTGT 273
Db 1 MetAlaAlaMetGluThrGluGlnLeuGlyValGluIlePheGluThrAlaAspCys 20
QY 274 GAGGAATATTGAATCAGAGATCGGCTTAATTTGAGGCTTTTATGTTGAGCGGTAT 333
Db 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
QY 334 TCCTGGAGTCAGCTTAAAAAGCTGCTTCCGATACCAAGAAATATCATGGCTACATGATG 393
Db 41 SerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
QY 394 GCTAAGGCACCATGATTTTCATGTTTGTGAAGAGGAATGATCCAGATGGACCTCATTC 453
Db 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
QY 454 GACAGAAATCTATTACCTTGCATGTCGCTGAGAACAGAGAAAATACACTGTTTATTCT 513
Db 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
QY 514 GAAATCCCAAAATCATTAATAGACAGCAGCTTAAATGCTCTCTGGAAGCCTCTTTTG 573
Db 101 GluIleProLysThrIleAsnArgAlaValLeuMetLeuSerTrpLysProLeuLeu 120
QY 574 GATCTTTTCAGGCACACTGGACTATGGAATGATTTCTCGAGAAGAAAGAACTATTAA 633
Db 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140
QY 634 GAAAGAAAACGCTTGAACAGCTCGGAATTCCTTACGATTATCACCAGGAAGTGA 693
Db 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160

QY 694 ACATTTCTGTTTCAAGCCGCTAGTGGAAATTTATCATGTAAGATGAGGGCCACAGGA 753
Db 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
QY 754 TTTAGCGCAACACCTTTTAAGCCCAATCTAGTGGAAACTAGTTGTCCTCCAACTACGGATG 813
Db 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
QY 814 GATCCAAAATTTATGCCCCGCTGATCCAGACTGAGTGTCTTTTATACATAGCAACGATATT 873
Db 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
QY 874 TGGATATCTAAACATCGTAAACACAGAGAAGAAAGAGAGCTCACATTATGTCACAATGAGCTA 933
Db 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisGlnGluLeu 240
QY 934 GCCAATCGAAGAAAGATGCCAGATCAGCTGGAGTGGCTACCTTTGCTTCTCCAAAGAGAA 993
Db 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu 260
QY 994 TTTGATAGATATTCTGGCTATTGGTGTGTCCTCAAGAGCTGAAACAACTCCCACTGGTGGT 1053
Db 261 PheAspArgTyrSerGlyTyrTrpCysProLysAlaGluThrThrProSerGlyGly 280
QY 1054 AAAATCTTAGAATTTCTATATGAGAAATGATGAATCTGAGCTGGAAATTTATTCATGTT 1113
Db 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300
QY 1114 ACATCCCCTATGTTGGAACAAGAGGGCAGATTCATTCGGTTATCTCTAAACAGGTACA 1173
Db 301 ThrSerProMetLeuGluThrArgAlaAspSerPheArgTyrProLysThrGlyThr 320
QY 1174 GCAATCTCTAAAGTCACATTTTAAGATGTCAGAAATATGATGATGCTGAGGAAGATC 1233
Db 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
QY 1234 ATAGATGTCATAGATAAGGAACCTAATTCACCTTTTTCAGATCTCTATTGGAAGAGTTGAA 1293
Db 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
QY 1294 TATATTGCCAGAGCTGGATGAGACTCTGAGGGAATAATATCTGGTCCATCTCATAGAT 1353
Db 361 TyrIleAlaArgAlaGlyTyrThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
QY 1354 CGCTCCCACTCGCTACAGATAGTGTGATCTCACCTGAGTCCCTGATTTCTGTGACGCCACTA 1413
Db 381 ArgSerGlnThrArgLeuGlnIleValLeuIleSerProGluLeuPheIleProValGlu 400
QY 1414 GATGATGTTATGSAAGGCAGAGACTCATTGAGTCAGTCCCTGATTTCTGTGACGCCACTA 1473
Db 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
QY 1474 ATTATCTATGAGAAGACACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTTTT 1533
Db 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
QY 1534 CCCCAGAGTCACGAGAGGAATTCAGTTTATTTTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1593
Db 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
QY 1594 CGTCATTTTACAAAATTCATCTATTTTAAAGGAACCAATATTAACGATCCAGTGGT 1653
Db 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
QY 1654 GGGCTGCTGCTCCCAAGTGAATTCAGTGTCTCTCAAGAGAGAGATGCAATTCACGAT 1713
Db 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 500
QY 1714 GGTGAATGGGAAGTCTTCTGGCCGATGATCTAATATCCAAAGTTGATGAAGTCAAGG 1773
Db 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520

```
QY 1774 CTGGTATATTTGAAGGCACCAAGACTCCCTTTAGAGCATCACCTGTACGTAGTCACT 1833
Db |||||||
QY 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuTyrValValSer 540
Db |||||||
QY 1834 TACGTAATCTCGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACATTTCTGCTGC 1893
Db |||||||
QY 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
Db |||||||
QY 1894 ATCAGTCAGCACTGTGACTTCTTTATAGTAAGTATAGTAACCCAGAGATCCACACTGT 1953
Db |||||||
QY 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
Db |||||||
QY 1954 GTGTCCTTTACAAGCTATCAAGTCCTGAGCATGACCCAACTTGCAAAACAAAGGAATTT 2013
Db |||||||
QY 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
Db |||||||
QY 2014 TGGGCCACCATTTTGGATTTCAGCAGGTCTCTTCTGACTATATCTCTCCAGAAATTTTC 2073
Db |||||||
QY 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProProGluIlePhe 620
Db |||||||
QY 2074 TCTTTTGAAGTACTACTGGATTACATTGTATGGGATGCTCTACAGCCCTCATGATCTA 2133
Db |||||||
QY 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrLysProHisAspLeu 640
Db |||||||
QY 2134 CAGCCTGGAAGAAATATCTACTGTGTCTGTTCATATATATGGTGGT 2178
Db |||||||
QY 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGly 655
Db |||||||
```

Search completed: April 14, 2006, 12:42:13
Job time : 527 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: April 14, 2006, 12:28:19 ; Search time 9.7 Seconds
(without alignments)
2733.328 Million cell updates/sec

Title: US-10-825-632-2
Perfect score: 5352
Sequence: 1 aadgtctaaagctccgagg.....agaattactaaaaaaaaa 3120

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 217505 seqs, 42489236 residues

Total number of hits satisfying chosen parameters: 435010

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-O=/abss/ABSSWEB spool/US10825632/runat 14042006 092402 3741/app_query.fasta_1
-DB=Published Applications AA New -OFMT=fastan -SUFFIX=n2p.rapbn -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DLOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0
-NAXLEN=2000000000 -HOST=abs02run
-USER=US10825632 @CGN 1 1 32 @runat 14042006 092402 3741 -NCPU=6 -ICPU=3
-NO_WMAP -NGS_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA New:

1: /SID85/ptodata/2/pubpaa/US08_NEW_PUB.pcp:*
2: /SID85/ptodata/2/pubpaa/US06_NEW_PUB.pcp:*
3: /SID85/ptodata/2/pubpaa/US07_NEW_PUB.pcp:*
4: /SID85/ptodata/2/pubpaa/PCT_NEW_PUB.pcp:*
5: /SID85/ptodata/2/pubpaa/US09_NEW_PUB.pcp:*
6: /SID85/ptodata/2/pubpaa/US10_NEW_PUB.pcp:*
7: /SID85/ptodata/2/pubpaa/US11_NEW_PUB.pcp:*
8: /SID85/ptodata/2/pubpaa/US60_NEW_PUB.pcp:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4700	84.7	882	7	US-11-151-601-20
2	651	11.7	745	7	US-11-079-463-6408
3	577	10.4	109	7	US-11-176-951-10
4	551	9.9	627	7	US-11-079-463-7758
5	530	9.5	762	7	US-11-116-939-13
6	529	9.5	738	7	US-11-208-288-4
7	529	9.5	766	6	US-10-501-035-234
8	529	9.5	766	6	US-11-208-288-2
9	525	9.5	766	6	US-10-522-789-2

10	517.5	9.3	760	7	US-11-208-288-6	Sequence 6, Appli
11	454	8.2	760	7	US-11-186-284-55	Sequence 55, Appl
12	394	7.1	99	7	US-11-176-951-11	Sequence 11, Appl
13	255	4.6	624	7	US-11-079-463-7504	Sequence 7504, Ap
14	249.5	4.5	657	7	US-11-179-977-1	Sequence 1, Appli
15	200.5	3.6	737	7	US-11-079-463-9281	Sequence 9281, Ap
16	145	2.6	102	7	US-11-176-951-7	Sequence 7, Appli
17	139	2.5	115	7	US-11-176-951-9	Sequence 9, Appli
18	133	2.5	115	7	US-11-176-951-12	Sequence 12, Appli
19	133.5	2.4	102	7	US-11-176-951-8	Sequence 8, Appli
20	128.5	2.3	1307	6	US-10-995-561-711	Sequence 711, App
21	125	2.3	24	7	US-11-176-951-16	Sequence 16, Appl
22	124.5	2.2	2483	7	US-11-186-999-2	Sequence 2, Appli
23	119.5	2.2	1019	6	US-10-995-561-982	Sequence 982, App
24	119	2.1	710	7	US-11-151-601-232	Sequence 23, Appl
25	118.5	2.1	1155	7	US-11-098-686-10550	Sequence 10550, A
26	118.5	2.1	1178	7	US-11-044-899-29	Sequence 29, Appl
27	118.5	2.1	1243	6	US-10-453-372-1136	Sequence 1136, Ap
28	118	2.1	668	6	US-10-454-437-118	Sequence 118, App
29	118	2.1	1243	6	US-10-453-372-1134	Sequence 1134, Ap
30	118	2.1	4913	6	US-10-453-372-1142	Sequence 1142, Ap
31	118	2.1	4961	6	US-10-453-372-1132	Sequence 1132, Ap
32	115	2.1	849	6	US-10-909-769-18	Sequence 18, Appl
33	113.5	2.0	420	7	US-11-079-463-6314	Sequence 6314, Ap
34	113.5	2.0	2458	7	US-11-186-999-11	Sequence 11, Appl
35	113.5	2.0	2515	7	US-11-113-424-53	Sequence 53, Appl
36	112.5	2.0	1285	7	US-11-206-071-2	Sequence 2, Appli
37	111.5	2.0	1206	6	US-10-995-561-709	Sequence 709, App
38	111.5	2.0	2256	7	US-11-144-368-4	Sequence 4, Appli
39	111.5	2.0	2455	7	US-11-186-999-4	Sequence 4, Appli
40	111.5	2.0	2458	7	US-11-186-999-6	Sequence 6, Appli
41	111.5	2.0	2458	7	US-11-186-999-13	Sequence 13, Appl
42	111	2.0	315	7	US-11-018-868-143	Sequence 143, App
43	110	2.0	1067	7	US/11/062	Sequence 3, Appli
44	110	2.0	1092	7	US/11/062	Sequence 6, Appli
45	110	2.0	2455	7	US-11-186-999-14	Sequence 14, Appli

ALIGNMENTS

RESULT 1

US-11-151-601-20
; Sequence 20 Application US/11/151601
; Publication No. US20060003413A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Olandt, Peter J.
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Curtiss, Rory A. J.
; APPLICANT: Williamson, Mark
; APPLICANT: Weich, Nadine
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASE, PHOSPHATASE,
; FILE OF INVENTION: AND PROTEASE FAMILY MEMBERS AND USES THEREOF
; CURRENT APPLICATION NUMBER: US/11/151,601
; CURRENT FILING DATE: 2005-06-13
; PRIOR APPLICATION NUMBER: US 10/170,789
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: US 09/797,039
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: PCT/US01/06525
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/186,061
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 09/882,166
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: PCT/US01/19269
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/212,078
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/934,406
; PRIOR FILING DATE: 2001-08-21


```
; PRIOR APPLICATION NUMBER: PCT/US01/26052
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/226,740
; PRIOR FILING DATE: 2000-08-21
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-151-601-20

Alignment Scores:
Pred. No.: 0 Length: 882
Score: 4700.00 Matches: 882
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 84.7% Indels: 0
DB: 7 Gaps: 0

US-10-825-632-2 (1-3120) x US-11-151-601-20 (1-882)
QY 214 ATGGCAGCAGCAATGGAACAGACAGCTGGGTGTTCAGATATTTCGAACTGCGGACTGT 273
DB 1 MetAlaAlaMetGlnThrGlnLeuGlnValGluLeuPheGlnThrAlaAspCys 20
QY 274 GAGGAGAAATATTGAATCAGAGATCGGCTTAATTTGGAGCCCTTTTATGTGAGCGGTAT 333
DB 21 GluGluAsnIleGluSerGlnAspArgProLysLeuGluProPheTyrValGluArgTyr 40
QY 334 TCCTGGAGTCAGCTAAAGAGCTGCTTGGCGATACAGAAATATATCATGCTACATGATG 393
DB 41 SerTrpSerGlnLeuLysLysLeuLeuAlaAspThrArgLysTyrHisGlyTyrMetMet 60
QY 394 GCTAAGGCACCATGATTTTCATGTTGTGAAGAGGAATGATCCAGATGGACCTCATTC 453
DB 61 AlaLysAlaProHisAspPheMetPheValLysArgAsnAspProAspGlyProHisSer 80
QY 454 GACAGAACTATTACCTTGGCATGCTGTGTGAGAACAGAGAAATACACTGTTTATTCT 513
DB 81 AspArgIleTyrTyrLeuAlaMetSerGlyGluAsnArgGluAsnThrLeuPheTyrSer 100
QY 514 GAAATTTCCAAACTATCAATAGACGACAGTCTTAATGCTCTCTTGGAGGCTCTTTTG 573
DB 101 GluIleProLysThrIleAsnArgAlaAlaValLeuMetLeuSerTrpLysProLeuLeu 120
QY 574 GATCTTTTCAGGCAACACTGGACTATGGAATGTATTCTCGAGAGAGAACTATTAAAGA 633
DB 121 AspLeuPheGlnAlaThrLeuAspTyrGlyMetTyrSerArgGluGluLeuLeuArg 140
QY 634 GAAAGAAACCATGGAAACAGTCGGNAATTGCTTCTTACGATTATCACCAGGAAGTGA 693
DB 141 GluArgLysArgIleGlyThrValGlyIleAlaSerTyrAspTyrHisGlnGlySerGly 160
QY 694 ACATTCTGTTTCAAGCCGGTAGTGGAAATTTATCAGTAAAGATGGAGGCGCCACAAGGA 753
DB 161 ThrPheLeuPheGlnAlaGlySerGlyIleTyrHisValLysAspGlyGlyProGlnGly 180
QY 754 TTTACGCAACACCTTTTAAGGCCCAATCTAGTCGAAACTAGTTGTCCCAACATACGGATG 813
DB 181 PheThrGlnGlnProLeuArgProAsnLeuValGluThrSerCysProAsnIleArgMet 200
QY 814 GATCCAAATATTATCCCGCTGTATCCAGACTGGATTGCTTTTATATACATAGCAACATATT 873
DB 201 AspProLysLeuCysProAlaAspProAspTrpIleAlaPheIleHisSerAsnAspIle 220
QY 874 TGGATATCTAATCATCGTAACAGAGAGAAAGAGACTCACTTATGTGCACATACAGCTA 933
DB 221 TrpIleSerAsnIleValThrArgGluGluArgLeuThrTyrValHisAsnGluLeu 240
QY 934 GCCAACATGGAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTTCTCCCAAGAA 993
DB
```

```
DB 241 AlaAsnMetGluGluAspAlaArgSerAlaGlyValAlaThrPheValLeuGlnGluGlu 260
QY 994 TTTGATAGATATTCTGGCTATTGCTGTCCCAAGAGCTGAAACAACTCCCACTGGTGTGT 1053
DB 261 PheAspArgTyrSerGlyTyrTrpCysProLysAlaGluThrThrProSerGlyGly 280
QY 1054 AAAATCTTAGAATTTCTATATGAAGAAATGATGAATCTGAGGTGGAAATATTATGTT 1113
DB 281 LysIleLeuArgIleLeuTyrGluGluAsnAspGluSerGluValGluIleIleHisVal 300
QY 1114 ACATCCCTATGTTGGAACCAAGAGGCGCAGATTCATTCCGTTATCTCTAAACACAGGTACA 1173
DB 301 ThrSerProMetLeuGluThrArgArgAlaAspSerPheArgTyrProLysThrGlyThr 320
QY 1174 GCAAAATCTTAAAGTCACTTTTAAAGATGTCAAGAAATAATGATTTGATGCTGAAGGAAGATC 1233
DB 321 AlaAsnProLysValThrPheLysMetSerGluIleMetIleAspAlaGluGlyArgIle 340
QY 1234 ATAGATGTCATAGATAAGCAACTAATTCACACTTTTTCAGATTCCTATTGGAAGAGATTGAA 1293
DB 341 IleAspValIleAspLysGluLeuIleGlnProPheGluIleLeuPheGluGlyValGlu 360
QY 1294 TATATTGCCAGAGCTGGATGGACTCCTGAGGGAATAATGCTTGGTCCATCTACTAGAT 1353
DB 361 TyrIleAlaArgAlaGlyTyrThrProGluGlyLysTyrAlaTrpSerIleLeuLeuAsp 380
QY 1354 CGCTCCCACTCGCTACAGATAGTGTGATCTCACCTGAATTTATTTATCCCACTAGAA 1413
DB 381 ArgSerGlnThrArgGluIleValLeuIleSerProGluLeuPheIleProValGlu 400
QY 1414 GATGATGTTATGGAAGCCAGACAGACTCATTTGAGTCAGTCGCTGATCTGTGAGCGCACTA 1473
DB 401 AspAspValMetGluArgGlnArgLeuIleGluSerValProAspSerValThrProLeu 420
QY 1474 ATTATCTATGAAGAAACCAACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTT 1533
DB 421 IleIleTyrGluGluThrThrAspIleTrpIleAsnIleHisAspIlePheHisValPhe 440
QY 1534 CCCCAGAGTCAGAGAGGAATAATGAGTTTATTTTCCCTCTGAATCCAAACACAGGTTTC 1593
DB 441 ProGlnSerHisGluGluGluIleGluPheIlePheAlaSerGluCysLysThrGlyPhe 460
QY 1594 CGTCAATTTATACAAATTTACATCTATTTTAAAGGAAACCAATAATAACATCCAGTGGT 1653
DB 461 ArgHisLeuTyrLysIleThrSerIleLeuLysGluSerLysTyrLysArgSerSerGly 480
QY 1654 GGGCTGCCTGCCCAAGTGATTTCAAGTGCTCTATCAAGAGGAGATAGCAATATACCAGT 1713
DB 481 GlyLeuProAlaProSerAspPheLysCysProIleLysGluGluIleAlaIleThrSer 500
QY 1714 GGTGAATGGGAAGTTCTTGGCGGCATGGATCTAATATCCAAAGTTGATGAAGTCAGAGG 1773
DB 501 GlyGluTrpGluValLeuGlyArgHisGlySerAsnIleGlnValAspGluValArgArg 520
QY 1774 CTGGTATATTTTGAAGGCACCAAGACTCCCTTTAGAGCATCATCTGTAGTGTAGTCACT 1833
DB 521 LeuValTyrPheGluGlyThrLysAspSerProLeuGluHisLeuLeuTyrValValSer 540
QY 1834 TAGTAAATCTCGAGAGAGTGACAGGCTGATCAGCGTGGCTACTACATCTCTGCTGTC 1893
DB 541 TyrValAsnProGlyGluValThrArgLeuThrAspArgGlyTyrSerHisSerCysCys 560
QY 1894 ATCAGTCAGCACTGTGACTCTTTATATAAGTAAAGTATATACACAGAGAAATCCACTGT 1953
DB 561 IleSerGlnHisCysAspPhePheIleSerLysTyrSerAsnGlnLysAsnProHisCys 580
QY 1954 GTGTCCCTTTTACAAGCTATCAAGTCTCGAAGATCACCACTTCGAAACCAAGGAATTT 2013
DB 581 ValSerLeuTyrLysLeuSerSerProGluAspAspProThrCysLysThrLysGluPhe 600
QY 2014 TGGGCCACCATTTTGGATTTCAGCAGGTCTCTTCTCTGACTATATCTCTCCCAAGAAATTT 2073
DB 601 TrpAlaThrIleLeuAspSerAlaGlyProLeuProAspTyrThrProGluIlePhe 620
```

```
QY 2074 TCTTTTGAAGTACTACTGGATTTCATTGTATGGGATGCTCTACAAAGCCTCATGATCTA 2133
Db 621 SerPheGluSerThrThrGlyPheThrLeuTyrGlyMetLeuTyrIysProHisAspLeu 640
QY 2134 CAGCCTGGAAAGAAATATCTACTGCTGTCTTCATATATGTTGCTCCTCAGGTGCGATTG 2193
Db 641 GlnProGlyLysLysTyrProThrValLeuPheIleTyrGlyGlyProGlnValGlnLeu 660
QY 2194 GTGAATATCCGTTTAAAGAGTCAAGTATTTCCGTTTGAATACCCCTAGCCTCTCAGGT 2253
Db 661 ValAsnAsnArgPheLysGlyValLysTyrPheArgLeuAsnThrLeuAlaSerLeuGly 680
QY 2254 TATGTGTTGTAGTATAGACAAACAGGGGATCCTGTCCAGCGGCTTAAATTTGAGGC 2313
Db 681 TyrValValValValIleAspAsnArgLysSerCysHisArgGlyLeuLysPheGluGly 700
QY 2314 GCCTTTAAATATAAATCGGTCAAAATAGAAATTCAGCATCAGGTGGAAGACTCCCAATAT 2373
Db 701 AlaPheLysTyrLysMetGlyGlnIleGluIleAspAspGlnValGluGlyLeuGlnTyr 720
QY 2374 CTAGCTCTCCGATATGATTTTCATTGACTTAGATCGTGTGGCATCCACGGCTGCTCTAT 2433
Db 721 LeuAlaSerArgTyrAspPheIleAspLeuAspArgValGlyIleHisGlyTyrSerTyr 740
QY 2434 GGAGGATACCTCTCCCTGATGGCATTAAATGCAGAGGTTCAGATATCTTCAGGGTGTCTAT 2493
Db 741 GlyGlyTyrLeuSerLeuMetAlaLeuMetGlnArgSerAspIlePheArgValAlaIle 760
QY 2494 GCTGGGGCCCATCTCTCTGTGGATCTTCTATGATACAGATACAGGACGTTATATG 2553
Db 761 AlaGlyAlaProValThrLeuTyrPheTyrAspThrGlyTyrThrGluArgTyrMet 780
QY 2554 GGTACCCCTGACCAAGTACAGCGCTATTTACTTAGATCTGTGGCATTCGACGAGCA 2613
Db 781 GlyHisProAspGlnAsnGluGlnGlyTyrTyrLeuGlySerValAlaMetGlnAlaGlu 800
QY 2614 AAGTTCCTCTCGAACCAATCGTTTACTGCTCTTACATGTTTCTCGATGAGATGTC 2673
Db 801 LysPheProSerGluProAsnArgLeuLeuLeuHisGlyPheLeuAspGluAsnVal 820
QY 2674 CATTTTGCACATACAGTATATTTACTGATTTTTTTAGTGGGCTGGAAAGCCATGAT 2733
Db 821 HisPheAlaHisThrSerIleLeuLeuSerPheLeuValArgAlaGlyLysProTyrAsp 840
QY 2734 TTACAGATCTATCTTCAGGAGACACAGACATAGAGTTCTCTGAATCGGAGAACATTAT 2793
Db 841 LeuGlnIleTyrProGlnGluArgHisSerIleArgValProGluSerGlyGluHisTyr 860
QY 2794 GAATGCACTCTTTTGCATCTACCTCAAGAAACCTTGGATCAGTATTTGCTGCTCTAAA 2853
Db 861 GluLeuHisLeuLeuHisTyrLeuGlnGluAsnLeuGlySerArgIleAlaLeuLys 880
QY 2854 GTGATA 2859
Db 881 ValIle 882
```

RESULT 2
US-11-079-463-6408
; Sequence 6408, Application US/11079463
; Publication No. US20060073161A1
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTERIOIDES FR
; FILE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: PATH00-03DIV2
; CURRENT APPLICATION NUMBER: US/11/079,463
; CURRENT FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/128,705
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: US 09/540,209
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444

```
; SEQ ID NO 6408
; LENGTH: 745
; TYPE: PRT
; ORGANISM: B.fragilis
US-11-079-463-6408

Alignment Scores:
Pred. No.: 1,14e-48 Length: 745
Score: 651.00 Matches: 199
Percent Similarity: 44.3% Conservative: 124
Best Local Similarity: 27.3% Mismatches: 244
Query Match: 11.7% Indels: 162
DB: 29

US-10-825-632-2 (1-3120) x US-11-079-463-6408 (1-745)
QY 730 GTAAAGATGGAGGCGCACAGGATTTAGCAACAACCTTTTAAGGCCCAATCTAGTGAA 789
Db 150 LeuSerAspGlyGlyProGln----- 156
QY 790 ACTAGTTGCCCCAACATACGATCGATCGATCCAAAATATATGCCCGGTGATCCGACTGGATT 849
Db 157 -----GlnAlaProValPheSerPro---AspGlyAsnLeuVal 168
QY 850 GCTTTTATACATAGCAACGATATTTGGATATCTAACATCGTAACCAAGAGAAAGGAGA 909
Db 169 AlaPheValArgAspAsnAsnIlePheLeuVal-----Lys 180
QY 910 CTCACTTATGTGCACAAATGAGCTAGCAACATCGGAAGAGATGCCAGA----- 957
Db 181 LeuLeuTyrGlyAsnSerGlu---SerGlnValThrGluAspGlyLysLeuAsnSerVal 199
QY 958 TCAGCTGGAGTCTGCTACCTTTGTCTCCAAAGAGAA-----TTTGATAGATATTTCTGGC 1011
Db 200 LeuAsnGlyIleProAspTyrValTyrGluGluGluPheGlyPheAsnArgAlaLeuGlu 219
QY 1012 TATTGTGTGTCTCAAAAGCTGAAACACTCCCACTGGTGTGTTAAATTTCTTAGAATTTCTA 1071
Db 220 Phe-----AsnAlaAspAsnThr-----MetLeuAla 228
QY 1072 TATGAGAAATAATGATGATCTGAGGTGGAAATATTTATGTTATCATCCCTATGTG--- 1128
Db 229 TyrValArgPheAspGluSerGluValPro-----SerTyrThrPheProLeuPheAla 246
QY 1129 ---GAAACAGGAGGCGCAGAT-----TCATTCCGTTAT 1158
Db 247 GlyGluAlaProArgTyrAspAlaLeuGlnAspTyrProGlyGluTyrThrTyrLysTyr 266
QY 1159 CCTAAACAGGTACAGCAATCTTAAGTC-----ACUTTTT----- 1194
Db 267 ProLysAlaGlyTyrProAsnSerLysValSerValHisThrPheAspIleLysSerLys 286
QY 1195 -----AAGATGTCAAAATATGATTGATGCTGAAGGAAGGATCATAGATGTCATAGAT 1248
Db 287 ValThrArgGlnValLysLeuProIleAspAlaAspGly----- 299
QY 1249 AAGGAACATTAATCAACCTTTTGTGATTTCTATTGAGGAGTGAATATATATTGCCAGAGCT 1308
Db 300 -----TyrIleProArgIle 304
QY 1309 GGATGAGCTCTCGAGGAAATAATGTTGTCCTACTAGATCGTCCCACTAGCTCCAGACTCGC 1368
Db 305 ArgPheThrGlnAspProAsnLysLeuAlaIleMetThrLeuAsnArgHisGlnAsnArg 324
QY 1369 CTACAGATAGTGTGATCTCACTGATTTATTTATCCAGTAGAAGATGATCTTATGGAA 1428
Db 325 PheAspMetTyrPheAlaAspPro----- 332
QY 1429 AGCAGAGACTCATTTGAGTCAGTCAGTCCTGATCTGTGACGCCCACTAATTTATCTATGAAGA 1488
Db 333 ArgSerThrValCysLysLeuAlaLeuArgAspGluSerProTyrTyrIleAsnGlu---- 351
QY 1489 ACAACAGACATCTGGATAAATAATCCATGACATCTTTTCATGATCTTTTCCCCCAAGTCACGAA 1548
```

Db	352	-----AsnValPheAspAenlleGlnPheTyProGlu-----	362	660	LysGluAenAlaGluGlyTyrLysAlaAlaSerAlaPheSerArgAlaAspAenLeu---	678
Qy	1549	GAGGAAATTCAGTTTATTTTGGCTCTGAATCGAAACAGGTTCCGTCATTATACAAA	1608	2623	TCTGAACCAATCGTTTACTCTCTTACATGGTTTCTCGATGAGAAATGTCATTTTGCA	2682
Db	363	-----TyrPheSerPheValSerAap-----LysSerGlyTyrProHisLeuTyr---	377	679	-----HisGlyAenLeuLeuValHisGlyMetAlaAspAenValHisPheGln	696
Qy	1609	ATTACATCTATTTTAAAGGAACAAATATAACGATCCAGTCGGTGGCTCCCTGCTCCA	1668	2683	CATACCATATATTACTAGTCTTTTACTGAGCTTTTACTGAGGCTGGAAGCCATATGATTTACAGATC	2742
Db	378	-----TTPtyrSerMetAenGlyAenLeu-----	385	697	AenCysThrGluTyrAlaGluHisLeuValGlnLeuGlyLysGlnPheAspMetGlnVal	716
Qy	1669	AGTGATTTCAAGTGTCTATCAAGAGGAGATAGCAATTTACAGTGGTGCAATGGAGATT	1728	2743	TATCTCTAGGAGACACAGCATTAAGAGTTCTCGATCCGGAGAACATTAAGACTGCAT	2802
Db	386	-----lleLysGln-----ValThrSerGlyAenTyrGluVal	396	717	TyrThrAenArAenHisSerIleTyrGlyAenThrArAenHisLeuTyrThrLys	736
Qy	1729	CTTGCGCGCATGATCTAATATCCAAAGTTGATGAAGTCAGAAAGCTGGTATATTTTGA	1788	2803	CTTTTGCACTACCTTCAAGAAAACCTT	2829
Db	397	LysAenPheIleGlyTyrAenProAspThrAenGlu-----PheTyrTyrThr	412	737	LeuThrAenPhePheArgAenAenLeu	745
Qy	1789	GGCACCAGACTCCCTTTAGAGCATCCTGTAGTACGTAGTACGTTAAATCCTGGA	1848	RESULT 3		
Db	413	SerAenGluSerProMetArgGlnAlaValTyrLysIleAap-----ArgLysGly	430	US-11-176-951-10		
Qy	1849	GAGTGACAAGGCTGACTGACCTGGCTACTCATCTTCTGTCATCAGTCAGCACTGT	1908	; Sequence 10, Application US/11176951		
Db	431	LysLysMetLysLeuSerAenGlnProGlyThrAenSerProIlePheSerSerSerMet	450	; Publication No. US20060024313A1		
Qy	1909	GACTTCTTTATAAGTATAGTAACCAAGAAATCCACACTGTGTCTCCCTTTACAAG	1968	; GENERAL INFORMATION:		
Db	451	LysTyrPheMetAenLysPheThrSerLeuAapThrProMetLeuIleThrLeu-----	468	; APPLICANT: CHEN, XIN		
Qy	1969	CTATCAAGTCTGAAAGATGACCCCACTTGCAAAAGAAATTTTGGGCCCACTTTTG	2028	; TITLE OF INVENTION: AGENTS THAT DISRUPT DIMER FORMATION IN DPP-IV FAMILY OF		
Db	469	-----AsnAapAenThrGlyLysValLeuLysThrLeuValThrAenAap	483	; FILE REFERENCE: 08842.0019		
Qy	2029	GATTCAGCAGTCTCTCTCTGACTATATCTCTCCA-----GAAATTTCTCTTTGAA	2082	; CURRENT APPLICATION NUMBER: US/11/176,951		
Db	484	LysLeuLysGlnLysLeuAlaGluTyrAlaIleProGlnLysGluPhePheThrPhelys	503	; CURRENT FILING DATE: 2005-07-06		
Qy	2083	AGTACTACTGGATTTACATTTGATGGATGCTCTACAAGCCTCATGATCTACAGCCTCGA	2142	; PRIOR FILING DATE: 2004-07-06		
Db	504	ThrThrGluGlyValAapLeuAenGlyTyrPheMetLysProValAenPheAapProAla	523	; PRIOR APPLICATION NUMBER: 60/586,095		
Qy	2143	AAGAAATATCTACTGTCTTATATATATGGTGGTCTCAGTGCGAGTTGGTGAATAAT	2202	; PRIOR FILING DATE: 2004-07-06		
Db	524	LysAsgTyrProValLeuMetPheGlnTyrSerGlyProGlySerGlnGlnValLeuAap	543	; PRIOR FILING DATE: 2004-07-06		
Qy	2203	CGGTTTAAAGAGTCAAGTATTTCCGCTTGAATACCTAGCCTCTCTAGTTATGTGGTT	2262	; NUMBER OF SEQ ID NOS: 22		
Db	544	LysTyrPheGlyIleSerTyrPheMetAlaSerLeuGlyTyrValVal	560	; SOFTWARE: Patentin Ver. 3.3		
Qy	2263	GTAGTGATAGACACAGGGGATCTGTACAGGGGCTTAAATTTCAAGCGCCTTTAAA	2322	; SEQ ID NO 10		
Db	561	AlaCysValAapGlyArgGlyThrGlyGlyArgGlySerGluPheGlnLysCysThrTyr	580	; LENGTH: 109		
Qy	2323	TATAAATGGGTCAATAGAAATTTGACGATCAGTGGAAGGACTCCCAATATCTAGCTTCT	2382	; ORGANISM: Homo sapiens		
Db	581	LeuAenLeuGlyValLysGluAlaLysAapGlnValGluAlaLysTyrLeuGlyGly	600	US-11-176-951-10		
Qy	2383	CGATATGATTTCAATGATCGTGTGGGATCCACGGCTGCTCTATGAGATAC	2442	Alignment Scores:		
Db	601	-----LeuProTyrValAapLysGlyArgGlyIleThrPheSerPheGlyGlyTyr	619	Pred. No.: 1,89e-42		
Qy	2443	CTCTCCCTGATGGCAATTAATCAGAGGTTCAGATATCTTACGGTGGTGTATTCGGGGCC	2502	Score: 577.00		
Db	620	MetThrIleMetSerMetSerMetSerGluGlyThrProValPheLysAlaGlyValAlaValAla	639	Percent Similarity: 100.0%		
Qy	2503	CCAGTCACTCTGTGGATCTTCTATGATACAGGATACACGAACTTATATGGGTCACTCT	2562	Best Local Similarity: 100.0%		
Db	640	AlaProThrAspTyrPheTyrAspThrValTyrThrGluArgPheMetArgThrPro	659	Query Match: 10.4%		
Qy	2563	GACCAAGTGAACAGGGCTATTTACTTAGGATCTTGTGGCCATGCAAGCAAGAAAGTTCCCC	2622	DB: 7		
Db	659	-----	678	US-10-825-632-2 (1-3120) x US-11-176-951-10 (1-109)		
Qy	2622	-----	696	Qy 2533 GGATACACGGAACGTTTATATGGTCCACCTGACCCAGAAATGAACAGGGCTATTACTTAGGA		
Db	678	-----	696	Db 1 GlyTyrThrGluArgTyrMetGlyHisProAspGlnAenGluGlnGlyTyrTyrLeuGly		
Qy	2692	-----	696	Qy 2593 TCTGTGGCCATGCAAGCAGAAAGTTCCCTCTGAACCAAAATCGTTTACTGCTCTTACAT		
Db	696	-----	696	Db 21 SerValAlaMetGlnAlaGluLysPheProSerGluProAenArgLeuLeuLeuHis		
Qy	2712	-----	696	Qy 2653 GGTTCCTCGATGAGAAATGTCATTTTGCATATCCATATATCTAGTATTTTACTGAGTTTGTAGTG		
Db	696	-----	696	Db 41 GlyPheLeuAapGluAenValHisPheAlaHisThrSerIleLeuLeuSerPheLeuVal		
Qy	2772	-----	696	Qy 2713 AGGCTCGGAAGCCATATGATTTTACAGATCTATCTCTCAGGAGAGACACAGCATTAAGATT		
Db	696	-----	696	Db 61 ArgAlaGlyLysProTyrAspLeuGlnIleTyrProGlnGluArgHisSerIleArgVal		
Qy	2832	-----	696	Qy 2773 CCTGAATCGGAGAAATTAATGATTCGATCTTTTGGCTACTCTTCAAGAAACCTTGA		
Db	696	-----	696	Db 81 ProGlnSerGlyGlnHisTyrGluLeuHisLeuLeuHisTyrLeuGlnGluAenLeuGly		
Qy	2859	-----	696	Qy 2833 TCAGGTATTCCTGCTCTAAAGTGATA		


```
Qy 1567 TTTCCTCTGAATGCAAAACAGGTTTCCTCAT-----TTATAC 1605
Db 349 TrpValGlyArgPheProSerGluProHisPheThrLeuAspGlyAsnSerPheTyr 368
Qy 1606 AAAATATACATCTATTAAAGGAAGCAAAATATAAAGATCAAGTGGTGGCTGCTGCT 1665
Db 369 LysIleIleSer-----AsnGluGluGlyTyrArgHisIle----- 380
Qy 1666 CCAAGTGAATTTCAAGTGTCTATCAAAGAGGAGATAGCAATATCCAGTGGTGAATGGAA 1725
Db 381 ---CyeTyrPheGlnIleAspLysIleAspCysThrPheIleThrLysGlyThrTrpGlu 399
Qy 1726 GTCTTGGCCGCATGATTAATCAAGTTCATGAAGTCAGAGAGCTGGTGTATATTTT 1785
Db 400 ValIleGly-----IleGluAlaLeuThrSerAspTyrLeuTyrIle 414
Qy 1786 GAAGGACC---AAAGACTCCCTTTAGAGCATCACCTGTAGCTAGCTAGCTAGTAAT 1842
Db 415 SerAsnGluTyrLysGlyMetProGlyGlyArgAsnLeuTyrIle 430
Qy 1843 CTGGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACATTTCTTGTGCATCAGT--- 1899
Db 431 -----GlnLeuSerAsp-----TyrThrLysValThrCysLeuSerCys 443
Qy 1900 -----CAGCAGCTGTGACTTCTTTATAAGTAAGTATAGTACACCAAGAATCCA 1947
Db 444 GluLeuAsnProGluArgCysGlnTyrTyrSerValSerPheSerLysGluAlaLysTyr 463
Qy 1948 CAC-----TGT-----GTGTCCTTTTCAAGCTATCAAGCTCTCAAGAT 1986
Db 464 TyrGlnLeuArgCysSerGlyProGlyLeuProLeuTyrThrLeuHisSerValAsn 483
Qy 1987 GACCCAACTTGCAAAACAAAGAAATTTGGGCCACCAATTTGGATTGAGAGTCTCTTT 2046
Db 484 AspLysGlyLeuArgValLeuGluAsp---AsnSerAlaLeuAspLysMet-----Leu 500
Qy 2047 CTGACTATCTCTCCAGAA-----ATTTCTCTTTTGAAGTACTACTGATTT 2097
Db 501 GlnAsnValGlnMetProSerLysLysLeuAspPheIleLeuAsnGluThrLysPhe 520
Qy 2098 ACATTTGTATGGATGCTCTACAGGCTCATGATCTACAGCTCGAAGAATATCTACT 2157
Db 521 ---TrpTyrGlnMetIleLeuProHis---PheAspLysSerLysTyrProLeu 538
Qy 2158 GTCTGTTTCATATATGGTCTCCTCAGTGCAGTTGGTGAATATCGTTTAAAGAGTC 2217
Db 539 LeuLeuAspValTyrAlaGlyProCysSerGln-----LysAlaAsp 552
Qy 2218 AAGTATTTCCGCTTGAAT-----ACCTAGCTCTCTAGGTTATGTGGTTGATGTG 2268
Db 553 ThrValPheArgLeuAsnTrpAlaThrTyrLeuAlaSerThrGluAsnIleIleValAla 572
Qy 2269 ---ATAGACAACAGGGGATCTCTGTCACCGAGGCTTAAATTTGAAGCGCTTTAAATAT 2325
Db 573 SerPheAspGlyArgGlySerGlyTyrGlnGlyAspLysIleMetHisAlaIleAsnArg 592
Qy 2326 AAAATGGGTCAATAGAAATGACATCAGTGGAGGACTCCAAATATCTAGCTTCTCGA 2385
Db 593 ArgLeuGlyThrPheGluValGluAspGlnIleGluAlaAlaArgGlnPhe---SerLys 611
Qy 2386 TATGATTTCAATGACTTAGATCGTGTGGCATCCACGGCTGGTCTATGAGGATACCTC 2445
Db 612 MetGlyPheValAspAsnLysArgIleAlaIleTrpGlyTrpSerTyrGlyTyrVal 631
Qy 2446 TCCTCATGCAATTAATGACAGGTCAGATATCTTCAGGGTGTCTATGCTGGGGCCCA 2505
Db 632 ThrSerMetValLeuGlySerGlySerGlyValPheLysCysGlyIleAlaValAlaPro 651
Qy 2506 GTCACCTGTGGATCTTCTATGATACAGGATACAGGAACGTTATATGGGT-----CAC 2559
Db 652 ValSerArgTrpGluTyrTyrAspSerValTyrThrGluArgTyrMetGlyLeuProThr 671
```

```
Qy 2560 CCTGACCAAGAATGAACAGGCTATTACTTAGATCTGTGGCCATCGCAAGCAAGAAAGTTTC 2619
Db 672 ProGluAsnLeuAspHisTyrArgAsnSerThrValMetSerArgAlaGluAsnPhe 691
Qy 2620 CCCTCTGAACCAAAATCGTTTACTGCTCTTACATGTTTCTCTGGATGAGAATGTCATTTT 2679
Db 692 -----LysGlnValGluTyrLeuLeuIleHisGlyThrAlaAspAspAsnValHisPhe 709
Qy 2680 GCACATACCATGATATTACTGAGTTTCTTACTGAGGCTGGAAAGCCATATGATTACAG 2739
Db 710 GlnGlnSerAlaGlnIleSerLysAlaLeuValAspValGlyValAspPheGlnAlaMet 729
Qy 2740 ATCTATCCTCCAGAGAGACACAGCATAAAGAGTTCCTGAATCGGAGACACATTATGAAC 2799
Db 730 TrpTyrThrAspGluAspHisGlyIleAlaSerSerThrAlaHisGlnHisIleTyrThr 749
Qy 2800 CATCTTTTGCATCTACCTTCAAGAA 2823
Db 750 HisMetSerHisPheIleLysGln 757
RESULT 6
US-11-208-288-4
; Sequence 4, Application US/11208288
; Publication No. US20060051366A1
; GENERAL INFORMATION:
; APPLICANT: CHANG, Chiwen
; TITLE OF INVENTION: USE OF SOLUBLE CD26 AS INHIBITOR OF
; FILE REFERENCE: 39533-0001
; CURRENT APPLICATION NUMBER: US/11/208,288
; PRIOR FILING DATE: 2005-08-18
; PRIOR APPLICATION NUMBER: US 60/605,013
; PRIOR FILING DATE: 2004-08-26
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 738
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-208-288-4
Alignment Scores:
Pred. No.: 5,66e-38 Length: 738
Score: 529.00 Matches: 186
Percent Similarity: 42.0% Conservative: 109
Best Local Similarity: 26.5% Mismatches: 280
Query Match: 9.5% Indels: 128
DB: 7 Gaps: 32
US-10-825-632-2 (1-3120) x US-11-208-288-4 (1-738)
Qy 829 CCCTGTGATCCAGACTGGATT-----GCTTTTATACAT 861
Db 121 ProAsnAsnThrGlnTrpValThrTrpSerProValGlyHisLysLeuAlaTyrValTrp 140
Qy 862 AGCAACCATATTTGGATATCTAACATCGTAACACAGAGAAGAAAGGAGACTCACTTATGTG 921
Db 141 AsnAsnAspIleTyrValLysIleGluProAsnLeuProSerTyrArgIleThrTrpThr 160
Qy 922 CACAATGAGTCAGCCAAACATGGAAGATGCCAGATCCAGATCGTGGAGTCGCTACCTTTGTT 981
Db 161 Gly-----LysGluAspIleIleTyrAsnGlyIleThrAspTrpVal 174
Qy 982 CTCACAAGAA-----TTTGATAGATATCTGCTGCTATTTGGTGTGTCCTCAAAAGCTGAAACA 1038
Db 175 TyrGluGluGluValPheSerAlaTyrSerAlaLeuTyrTrpSerProAsnGlyThrPhe 194
Qy 1039 ACTCCAGTGGTGAATAATCTTAGAATCTTATGAGAAATGATGAATCTGAGGTG 1098
Db 195 -----LeuAlaTyrAlaGlnPheAsnAspThrGluVal 205
Qy 1099 GAAATTATTCAT-----GTTACATCCCTATGTTGGAACAAGAGGGCGAGATCA 1149
```


206 ProLeuIleGluTyrSerPheTyrSerAspGluSerLeuGlnTyrProLysThrValArg 225
1150 TTCGGTTATCCTAAACACAGGTACAGCAAAATCCTAAAGTCACATTTTAAAGATGTCAGAAATA 1209
226 ValProTyrProLysAlaGlyAlaValAsnProThrValLysPhePheVal-----Val 243
1210 ATGATTGATCCTGAAGAGGATCATAGATGTCATAGATAAGGAACCTAAATCAACCTTTT 1269
244 AsnThrAspSerLeuSerValThrAsnAlaThrSerIleGlnIleThrAlaProAla 263
1270 GAGATTCTATTGACGAGTTGAATATATTCACAGACTCCGCTACACATAGTGTTCATCTCA 1389
264 SerMetLeuIle---GlyAspHisTyrLeuLysCysAspValThrTrpAla----- 278
1330 TATGCTGGTCCATCTACTAGATCGCTCCGAGACTCCGCTACACATAGTGTTCATCTCA 1389
279 -----ThrGlnGluArgIleSerLeuGlnTrpLeu--- 288
1390 CCTGAATTATTCAGTAGAAGATGATTATGGAAAGCAGACAGACTCATTTGAGTCA 1449
289 -----ArgArgIleGlnAsn 293
1450 GTGCTGATTCGTGACGCCACTAATATCTATGAAGAAACACAGACATCTGGATTAAT 1509
294 Tyr-----SerValMetAspIleCysAspTyrAspGluSerSerGlyArgTrp---Asn 310
1510 ATCCATGACATCTTTTCATGTTTTCCTCCAAAGTCACGAAAGAGGAAATTCAGTTTATTTT 1569
311 Cys-----LeuValAlaArgGlnHisIleGluMetSerThrThrGlyTrp 325
1570 GCCTCTGAATCAAAACAGGTTTCCGTCAT-----TTATACAAA 1608
326 ValGlyArgPheAspGluProHisPheThrLeuAspGlyAsnSerPheTyrLys 345
1609 ATTACATCTATTTAAAGGAAGCAAAATATAAACGATCCAGTGGTGGCTGCTCTCCA 1668
346 IleIleSer-----AsnGluGluGlyTyrArgHisIle----- 356
1669 AGTGATTTCAAGTCTCTATCAAGAGGAGATAGCAATACAGTGGTGAATGGGAAGTT 1728
357 CysTyrPheGlnIleAspLysLysCysThrPheIleThrLysGlyThrTrpGluVal 376
1729 CTGGCCGGCATGATCTAATATCAAGTTGTATGATGAGTCAGAGCTGGTATATTTGAA 1788
377 IleGly-----IleGluAlaLeuThrSerAspTyrLeuTyrTyrIleSer 391
1789 GGCACC---AAAGACTCCCTTTAGAGCATCCTGTGATCGTAGTCAGTTACGTAATCCT 1845
392 AsnGluTyrLysGlyMetProGlyArgAsnLeuTyrLysIleGlnLeuIleAspTyr 411
1846 GGAGAGGTGACAAAGCTGACGCGGTGCTACTACATCTTGTGTCATCAGT---CAG 1902
412 ThrLysValThrCysLeu-----SerCysGluLeuAsnProGlu 424
1903 CACTGTGACTCTTTTATAGTAAGTATAGTAACAGAAATCCACAC-----TGT 1953
425 ArgCysGlnTyrTyrSerValSerPheSerLysGluAlaLysTyrTyrGlnLeuArgCys 444
1954 -----GTGCTCCTTTACAGCTATCAAGTCTCCTGAAGATGACCCCACTTGCAA 2001
445 SerGlyProGlyLeuProLeuTyrThrLeuHisSerSerValAsnAspLysGlyLeuArg 464
2002 ACAAGGAATTTGGCCACCATTTTGGATTTCAGAGTCTCTCTCTGACTATATCTCCT 2061
465 ValLeuGluAsp---AsnSerAlaLeuAspLysMet-----LeuGlnAsnValGlnMet 481
2062 CAGAA-----ATTTCTCTTTGAAAGTACTACTGGATTACATTTGATGGATG 2112
482 ProSerLysLeuAspPheIleIleLeuAsnGlnThrLysPhe---TrpTyrGlnMet 500
2113 CTCTACAGCTCATGATCTACAGCTCGAAGAAATATCTACTGCTGCTGTTATATAT 2172
501 IleLeuProHis-----PheAspLysSerLysLysTyrProLeuLeuAspValTyr 519

2173 GGTGGTCTCAGGTGCGAGTTGGTGATATATCGGTTTAAAGGAGTCAAGTATTTTCGGCTTG 2232
520 AlaGlyProCysSerGln-----LysAlaAspThrValPheArgLeu 533
2233 AAT-----ACCTAGCCTCTCTAGGTTATGGGTGTAGTG---ATAGACAACAGG 2280
534 AsnTrpAlaThrTyrLeuAlaSerThrGluAsnIleIleValAlaSerPheAspGlyArg 553
2281 CGATCTGTTCACCGAGGCTTAAATTTGAAGGCGCTTTAAATATATAAATGGTCAATA 2340
554 GlySerGlyTyrGlnGlyAspLysIleMetHisAlaIleAsnArgArgLeuGlyThrPhe 573
2341 GAAATTCACCATCAGGTGGAAGGACTCCCAATATCTAGCTTCTCGATATGATTTTCAATGAC 2400
574 GluValGluAspGlnIleGluAlaAlaArgGlnPhe---SerLysMetGlyPheValAsp 592
2401 TTAGATCGTGTGGGCATCCACGCTGCTCTATGGAGGATACCTCTCCCTGATGCAATTA 2460
593 AsnLysArgIleAlaIleTrpGlyTyrSerTyrGlyTyrValThrSerMetValLeu 612
2461 ATGCAGAGGTCCAGATATCTTCAGGTTGCTTATGCTGGGGCCCCAGTCACTCTGTGGATC 2520
613 GlySerGlySerGlyValPheLysCysGlyIleAlaValAlaProValSerArgTrpGlu 632
2521 TTCTATGATACAGGATACAGGACGTTATATCGGT-----CACCTCACCAGAAATGA 2574
633 TyrTyrAspSerValTyrThrGluArgTyrMetGlyLeuProThrProGluAspAsnLeu 652
2575 CAGGCTATTACTTAGGATCTGTGGCCATCGAAGCAGAAAGTTCCCTCTGAACCAAAAT 2634
653 AspHisTyrArgAsnSerThrValMetSerArgAlaGluAsnPhe-----LysGlnVal 670
2635 CTTTTACTGCTCTACATGTTCTCGATCAGAAATGTCATTTTCACATACCATGATATA 2694
671 GluTyrLeuLeuIleHisGlyThrAlaAspAspAsnValHisPheGlnGlnSerAlaGln 690
2695 TTACTGAGTTTTTTAGTGAGGCTGGAAGCCATATGATTACAGATCTATCTCAGAG 2754
691 IleSerLysAlaLeuValAspValGlyValAspPheGlnAlaMetTrpTyrThrAspGlu 710
2755 AGACACAGCATAAAGAGTTCTCGAATCGGAGAACATATGAACTGGCATCTTTTGGCACTAC 2814
711 AspHisGlyIleAlaSerSerThrAlaHisGlnHisIleTyrThrHisMetSerHisPhe 730
2815 CTTCAAGAA 2823
731 IleLysGln 733
RESULT 7
US-10-501-035-234
; Sequence 234, Application US/10501035
; Publication No. US20060046249A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: IDENTIFICATION OF POLYNUCLEOTIDES AND POLYPEPTIDE FOR PREDICTING
; TITLE OF INVENTION: ACTIVITY OF COMPOUNDS THAT INTERACT WITH PROTEIN TYROSINE KINASE
; TITLE OF INVENTION: AND/OR PROTEIN TYROSINE KINASE PATHWAYS
; FILE REFERENCES: D0185 PCT
; CURRENT FILING DATE: 2004-07-09
; PRIOR FILING DATE: 2002-01-18
; NUMBER OF SEQ ID NOS: 795
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 234
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-501-035-234
Alignment Scores: 5.72e-38 Length: 766
Pred. No.:

Score: 529.00 Matches: 186
Percent Similarity: 42.0% Conservative: 109
Best Local Similarity: 26.5% Mismatches: 280
Query Match: 9.5% Indels: 128
DB: 6 Gaps: 32

US-10-825-632-2 (1-3120) x US-10-501-035-234 (1-766)

QY 829 CCGCTCATCCAGCTGGATT-----GCTTTATACAT 861
DB 149 ProAsnThrGlnTrpValThrTrpSerProValGlyHisLysLeuAlaTyrValTrp 168
QY 862 AGCAACGATATTGGATATCAATCGTAACGAGAGAGAAAGAGAGACTCATATATGG 921
DB 169 AsnAsnAspIleTyrValLysLysGluProAsnLeuProSerTyrArgIleThrTrpThr 188
QY 922 CACATGAGCTAGCAACATGGAAGAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTT 981
DB 189 Gly-----LysGluAspIleIleTyrAsnGlyIleThrAspTrpVal 202
QY 982 CTCCAAAGAA--TTTGATAGATATTCTGGCTATTGGTGTGTGCCAAAGAGCTGAAACA 1038
DB 203 TyrGluGluGluValPheSerAlaTyrSerAlaLeuTrpTrpSerProAsnGlyThrPhe 222
QY 1039 ACTCCAGTGGTGTAAATTTCTAGAAATTTCTATGAAATAATGATGAATCTCAGGTG 1098
DB 223 -----LeuAlaTyrAlaGlnPheAsnAspThrGluVal 233
QY 1099 GAAATTTATCAT-----GTTACATCCCTATGTTGGAAACAGAGAGGCGCAGATTCA 1149
DB 234 ProLeuIleGluTyrSerPheTyrSerAspGluSerLeuGlnTyrProLysThrValArg 253
QY 1150 TTCGGTTATCTTAAACAGCTACAGCAATCTTAAAGTCACATTTTAAAGATGTCAGAAATA 1209
DB 254 ValProTyrProLysAlaGlyAlaValAsnProThrValLysPheVal-----Val 271
QY 1210 ATGATTTCATGCTGAAGAGATCATAGATGTCATAGATAAGAACTAATTCACCTTTT 1269
DB 272 AsnThrAspSerLeuSerValThrAsnAlaThrSerIleGlnIleThrAlaProAla 291
QY 1270 GAGATTCTATTGAAGAGTGAATATATTGCGAGAGCTGGATGGATCCTCGAGGAAAA 1329
DB 292 SerMetLeuIle---GlyAspHisTyrLeuCysAspValThrTrpAla----- 306
QY 1330 TATGCTGGTCCATCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCA 1389
DB 307 -----ThrGlnGluArgIleSerLeuGlnTrpLeu--- 316
QY 1390 CCTGAATTTATTTATCCAGTAGAAGATGATGTTATGGAAGGCGAGAGACTCATAGTCA 1449
DB 317 -----ArgArgIleGlnAsn 321
QY 1450 GTGCTGATTTCTGACGCCCACTAATTTATCTATGAAGAAACAAACAGACATCTGGTAAT 1509
DB 322 Tyr-----SerValMetAspIleCysAspTyrAspGluSerSerGlyArgTrp---Asn 338
QY 1510 ATCCATGACATCTTTTCATGTTTTTCCCAAGTACGAGAGAGAAATGAGTTATTTT 1569
DB 339 Cys-----LeuValAlaArgGlnHisIleGluMetSerThrThrGlyTrp 353
QY 1570 GCCTCTGAATGCAAAACAGGTTTCCGTCAT-----TTATACAAA 1608
DB 354 ValGlyArgPheArgProSerGluProPheThrLeuAspGlyAsnSerPheTyrLys 373
QY 1609 ATTACATCTATTTTAAAGGAAAGCAAAATATAACGATCCAGTGGTGGCTGCTGCCA 1668
DB 374 IleIleSer-----AsnGluGluGlyTyrArgHisIle----- 384
QY 1669 AGTGATTTCAAGTGTCTTATCAAGAGAGAGATAGCAATTTACCAAGTGGTGAATGGGAGTT 1728
DB 385 CysTyrPheGlnIleAspLysAspCysThrPheIleThrLysGlyThrTrpGluVal 404
QY 1729 CTTGGCCGCGATGATCTAATATCCAAAGTTGATGAACTCAGAGGCTGGTATTTTGA 1788

DB 405 IleGly-----IleGluAlaLeuThrSerAspTyrLeuTyrTyrIleSer 419
QY 1789 GGCACC---AAAGACTCCCTTTAGAGCATCACCTGTACGTAGTACGTACGTAAATCTCT 1845
DB 420 AsnGluTyrLysGlyMetProGlyGlyArgAsnLeuTyrLysIleGlnLeuIleAspTyr 439
QY 1846 GGAGAGGTGACAGGCTGACTGACGGTGGCTACTACATTTCTTCTGCTGCATCAGT---CAG 1902
DB 440 ThrLysValThrCysLeu-----SerCysGluLeuAsnProGlu 452
QY 1903 CACTGTGACTCTTTTATAAGTAGTAGTAACACAGAGAATCCACAC-----TGT 1953
DB 453 ArgCysGlnTyrTyrSerValSerPheSerLysGluAlaLysTyrTyrGlnLeuArgCys 472
QY 1954 -----GTGCTCCCTTTTCAAGCTATCAAGTCTCTCAAGATGACCCCACTTGCAAA 2001
DB 473 SerGlyProGlyLeuProLeuTyrThrLeuHisSerSerValAsnAspLysGlyLeuArg 492
QY 2002 ACAAGGAATTTTGGGCCACCATTTTGGATTGAGTTCAGAGGTCTCTTCTGCTGACTACTCT 2061
DB 493 ValLeuGluAsp---AsnSerAlaLeuAspLysMet-----LeuGlnAsnValGlnMet 509
QY 2062 CCAGAA-----ATTTCTCTTTTGAAGTACTACTGATTTTACATTTGATGGGATG 2112
DB 510 ProSerLysLysLeuAspPheIleLeuAsnGluThrLysPhe---TrpTyrGlnMet 528
QY 2113 CTCCTCAAGCTCATGATCTACAGCTCGAAAGAAATATCTCTACTGTCTGTTCATATAT 2172
DB 529 IleLeuProPheHis---PheAspLysSerLysLysTyrProLeuLeuLeuAspValTyr 547
QY 2173 GTGTGCTCTCAGGTGAGTGGTGAATATCGTTTAAAGAGTCAAGTATTTCGGCTTG 2232
DB 548 AlaGlyProCysSerGln-----LysAlaAspThrValPheArgLeu 561
QY 2233 AAT-----ACCTAGCTCTCTAGTATTATGTTGGTTAGTG---ATAGACAAACAGG 2280
DB 562 AsnTrpAlaThrTyrLeuAlaSerThrGluAsnIleIleValAlaSerPheAspGlyArg 581
QY 2281 GGATCTGTCGACCGAGGCTTAAATTTGAAGCGCTTTTAAATATAAATCGGTCAATA 2340
DB 582 GlySerGlyTyrGlnGlyAspLysIleMethIleAlaIleAsnArgArgLeuGlyThrPhe 601
QY 2341 GAAATTCAGCATCAGGTGGAAGGATCCCAATATCTAGCTTCTCGATATGATTTCATTCAC 2400
DB 602 GluValGluAspGlnIleGluAlaAlaArgGlnPhe---SerLysMetGlyPheValAsp 620
QY 2401 TTAGATCGTGGGATCCACCGCTGCTCTTATGAGGATACCTCTCCCTCATGCGCATTA 2460
DB 621 AsnLysArgIleAlaIleTrpGlyTrpSerGlyTyrValThrSerMetValLeu 640
QY 2461 ATGACAGGTTCAGATATCTTCAAGGTGCTATTGTGGGCGCCAGTCACTCTGTGGATC 2520
DB 641 GlySerGlySerGlyValPheLysCysGlyIleAlaValAlaProValSerArgTrpGlu 660
QY 2521 TTCTATCATCAGATGATACACGGAAGCTTATATGGT-----CACCTGACAGAAATA 2574
DB 661 TyrTyrAspSerValTyrThrGluArgTyrMetGlyLeuProThrProGluAspAsnLeu 680
QY 2575 CAGGGCTATTACTTAGGATCTGTGGCATGCGAGAGAGAAAGTTCCTCTGGAACCAAT 2634
DB 681 AspHisTyrArgAsnSerThrValMetSerArgAlaGluAsnPhe-----LysGlnVal 698
QY 2635 CGTTTACTGCTCTTACATGTTTCTCGATGAGATGATGTCATTTTGCACATACAGTATA 2694
DB 699 GluTyrLeuLeuIleHisGlyThrAlaAspAsnValHisPheGlnSerAlaGln 718
QY 2695 TTACTAGTTTCTTACTGAGGGCTGAAAGCCATATGATTACAGATCTATCTCTCAGAG 2754
DB 719 IleSerLysAlaLeuValAspValGlyValAspPheGlnAlaMetTrpTyrThrAspGlu 738
QY 2755 AGACACAGCATAGAGTTCTGTAATCGGAGAACATTTATGAATCTGATCTTTTGCATC 2814


```
Db 602 GluValGluAspGlnIleGluAlaAlaArgGlnPhe---SerLysMetGlyPheValAsp 620
QY 2401 TTAGATCGTGTGGGATCCACGGCTGCTCTATGAGGATACCTCTCCCTGATGGCATTA 2460
Db 621 AsnLysArgIleAlaIleTrpGlyTrpSerTyrGlyGlyTyrValThrSerMetValLeu 640
QY 2461 ATGCAGAGGTGAGATATCTTCAGGGTTGCTATGCTGGGGCCCGAGTCACTCTGTGGATC 2520
Db 641 GlySerGlySerGlyValPheLysCysGlyIleAlaValAlaProValSerArgTrpGlu 660
QY 2521 TTCATGATACAGGATACACGGAACGTTATATGGGT-----CACCTGCACGAGATGAA 2574
Db 661 TyrTyrAspSerValTyrThrGluArgTyrMetGlyLeuProThrProGluAspAsnLeu 680
QY 2575 CAGGGCTATTACTTAGATCTGTGGCCATGCACAGCAGAAAGTTCCCTCTGAACCAAT 2634
Db 681 AspHisTyrArgAsnSerThrValMetSerArgAlaGluAsnPhe-----LysGlnVal 698
QY 2635 CGTTTACTGCTTACATGTTTCCCTGGATGAGATGTCCATTTGACATACCATGATATA 2694
Db 699 GluTyrLeuLeuIleHisGlyThrAlaAspAspAsnValHisPheGlnGlnSerAlaGln 718
QY 2695 TTACTGAGTTTTTTAGTGGGGCTGGAAGCCATATGATTTACAGATCTATCCTCAGGAG 2754
Db 719 IleSerLysAlaLeuValAspValGlyValAspPheGlnAlaMetTrpTyrThrAspGlu 738
QY 2755 AGACACACATAGAGTTCTCGAATCGGAGACATATGAACTGCATCTTTTGGCACTAC 2814
Db 739 AspHisGlyIleAlaSerSerThrAlaHisGlnHisIleTyrThrHisMetSerHisPhe 758
QY 2815 CTTCAAGAA 2823
Db 759 IleLysGln 761
```

RESULT 9

```
US-10-522-789-2
; Sequence 2, Application US/10522789
; Publication No. US20050260732A1
; GENERAL INFORMATION:
; APPLICANT: TANABE SEIYAKU CO., LTD.
; TITLE OF INVENTION: Three-dimensional structure of dipeptidyl peptidase IV
; FILE REFERENCE: 03-039-PCT
; CURRENT APPLICATION NUMBER: US/10/522,789
; CURRENT FILING DATE: 2005-01-28
; PRIOR APPLICATION NUMBER: US 60/398,761
; PRIOR FILING DATE: 2002-07-29
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 2
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-522-789-2
```

Alignment Scores:

Pred. No.:	1-28e-37	Length:	766
Score:	525.00	Matches:	186
Percent Similarity:	41.6%	Conservative:	108
Best Local Similarity:	26.3%	Mismatches:	277
Query Match:	9.5%	Indels:	136
DB:	6	Gaps:	33

US-10-825-632-2 (1-3120) x US-10-522-789-2 (1-766)

```
QY 829 CCGCTGATCCAGACTGGATT-----GCTTTTATACAT 861
Db 149 ProAsnAsnThrGlnTrpValThrTrpSerProValGlyHisLysLeuAlaTyrValTrp 168
QY 862 AGCAACGATATTGATGATCTTAACATCGTAACACGAGAGAAAGGAGACTCATCTATGTC 921
Db 169 AsnAsnAspIleTyrValLysIleGluProAsnLeuProSerTyrArgIleThrTrpThr 188
```

```
QY 922 CACAATGAGCTAGCCAAACATGGAAGAAGATGCCAGATCAGCTGGAGTCGCTACCTTTGTT 981
Db 189 Gly-----LysGluAspIleIleTyrAsnGlyIleThrAspTrpVal 202
QY 982 CTCCAAGAAGAA---TTTGATAGATATTTCTGGCTATTTGGTGTGTCACAAAGAGCTGAACA 1038
Db 203 TyrGluGluGluValPheSerAlaTyrSerAlaLeuTrpTrpSerProAsnGlyThrPhe 222
QY 1039 ACTCCCGAGTGGTGAATAATCTTAGAATTCATATGAAGAAAATGATGAATCTGAGGTG 1098
Db 223 -----LeuAlaTyrAlaGlnPheAsnAspThrGluVal 233
QY 1099 GAAATTAATTCAT-----GTTACATCCCCCTATGTTGGAACAAGAGGAGCGAGATTCA 1149
Db 234 ProLeuIleGluTyrSerPheTyrSerAspGluSerLeuGlnTyrProLysThrValArg 253
QY 1150 TTCGGTTATCTAAAAACAGGTACAGCAATCCTAAAGTCACCTTTTAAGATCTCGAATA 1209
Db 254 ValProTyrProLysAlaGlyAlaValAsnProThrValLysPhePheVal-----Val 271
QY 1210 ATGATTGATGCTGGAAGGAAGATCATAGATGTCATAGATAGGAACAACTAATTCACACCTTT 1269
Db 272 AsnThrAspSerLeuSerValThrAsnAlaThrSerIleGlnIleThrAlaProAla 291
QY 1270 GAGATTCTATTGAAGGAGTTGNAATATATATGTCAGAGCTGGATGGAACCTCTCTGAGGAAA 1329
Db 292 SerMetLeuIle---GlyAspHisTyrLeuCysAspValThrTrpAla----- 306
QY 1330 TATGCTTGGTCCATCTACTAGATCGCTCCAGACTCGCTACAGATAGTGTGATCTCA 1389
Db 307 -----ThrGlnGluArgIleSerLeuGlnTrpLeu--- 316
QY 1390 CCTGAATTAATTTATCCAGTAGAAGATGATGTTATGGAAGGACAGAGACTCATTTAGTCA 1449
Db 317 -----ArgArgIleGlnAsn 321
QY 1450 GTGCTGATTTCTGTGACGCCACTAATATCTATGATGAAGAAAACAACAGACATCTGGATAAAT 1509
Db 322 Tyr-----SerValMetAspIleCysAspTyrAspGluSerSerGlyArgTrp---Asn 338
QY 1510 ATCCATGACATCTTTCAATGTTTTCCTCCCAAGTCTCAGAGAGGAAATTTAGTTATTTT 1569
Db 339 Cys-----LeuValAlaArgGlnHisIleGluMetSerThrThrGlyTrp 353
QY 1570 GCCTCTGAATGCAAAACAGTTTCCGTCA-----TTATACAAA 1608
Db 354 ValGlyArgPheArgProSerGluProHisPheThrLeuAspGlyAsnSerPheTyrLys 373
QY 1609 ATTACATCTATTTTAAAGGAAGCAAAATATAACAGTCCAGTGGTGGCTGCTCTCCA 1668
Db 374 IleIleSer-----AsnGluGluGlyTyrArgHisIle----- 384
QY 1669 AGTGATTTCAAGTGTCTTATCAAGAGAGAGATAGCAATTAACAGTGGTGAATGGGAAGTT 1728
Db 385 CysTyrPheGlnIleAspLysLysAspCysThrPheIleThrLysGlyThrTrpGluVal 404
QY 1729 CTTGGCCGGCATGGATCTAATATPCCAAAGTTGATGAAGTCAAGAGCTGTATATTTGAA 1788
Db 405 IleGly-----IleGluAlaLeuThrSerAspTyrLeuTyrTyrIleSer 419
QY 1789 GGCACC---AAAGACTCCCCCTTTAGAGCATCACTGTAGTACGTACGTACGTAATCTCT 1845
Db 420 AsnGluTyrLysGlyMetProGlyGlyArgAsnLeuTyrLysIle----- 434
QY 1846 GGAGAGGTGACAAGGCTGACTGACCTGGCTACTACATTTCTTGTGTCATCAGT----- 1899
Db 435 -----GlnLeuSerAsp-----TyrThrLysValThrCysLeuSerCysGlu 448
QY 1900 -----CAGCACTGTGACTTCTTTTATAAGTAAAGTAACTAGTAAACAGAGAATCCACAC 1950
Db 449 LeuAsnProGluArgCysGlnTyrTyrSerValSerPheSerLysGluAlaLysTyrTyr 468
QY 1951 -----TGT-----GTGTCCTTTTACAAGCTATCAAGTCTCTGAAGATGAC 1989
```

```
Db 469 GlnLeuArgCysSerGlyProGlyLeuProLeuTyrThrLeuHisSerValAsnAsp 488
; CURRENT FILING DATE: 2005-08-18
; PRIOR APPLICATION NUMBER: US 60/605,013
; PRIOR FILING DATE: 2004-08-26
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 760
; TYPE: PRT
; ORGANISM: Mus musculus
; US-11-208-288-6
Qy 1990 CCAACTGCAAAACAAAGGAATTTGGCCACCAATTTGGATTTCAGCAGGCTCTTCTCT 2049
Db 489 LysGlyLeuArgValLeuGluAsp---AsnSerAlaLeuAspLysMet-----LeuGln 505
Qy 2050 GACTATATCTCTCCAGAA-----ATTTCCTCTTTGAAAGACTACTACTGGATTACA 2100
Db 506 AsnValGlnMetProSerLysLysLeuAspPheLeileLeuAsnGlnThrLysPhe--- 524
Qy 2101 TTGTATGGATGCTCTCAAGCCTCATGATCTACGCCCTGGAAAGAAATATCTACTGTG 2160
Db 525 TrpTyrGlnMetIleLeuProPheHis---PheAspLysSerLysLysTyrProLeuLeu 543
Qy 2161 CTGTTCATATATGTGTCTCTCAGGTCGAGTGTGGTAATAATCGGTTTAAAGGAGCTCAAG 2220
Db 544 LeuAspValTyrAlaGlyProCysSerGln-----LysAlaAspThr 557
Qy 2221 TATTTCGCTTGAAT-----ACCTAGCTCTCTAGTTATGTGTGTGTAGTG--- 2268
Db 558 ValPheArgLeuAsnTyrPalatThrTyrLeuAlaSerThrGluAsnIleileValAlaSer 577
Qy 2269 ATAGACAACAGGGATCTGTCCACGAGGCTTAAATTTGAAGGCGCTTTAAATATAA 2328
Db 578 PheAspGlyArgGlySerGlyTyrGlnGlyAspLysIleMetHisAlaIleAsnArgArg 597
Qy 2329 ATGGGTCAAATAGAAATTCAGATCAGGTGGAGGACTCCAATATCTAGCTTCTCGATAT 2388
Db 598 LeuGlyThrPheGluValGluAspGlnIleGluAlaArgGlnPhe---SerLysMet 616
Qy 2389 GATTTCATGCTTATAGATCTGTGGGCTCCAGGCTCCAGGCTCTTATGGAGGATACCTCC 2448
Db 617 GlyPheValAsnLysArgIleAlaIleTrpGlyTyrSerTyrGlyTyrValThr 636
Qy 2449 CTGATGGCAATTAATCAGAGGTGATATCTTCAGGTTGCTATGCTGGGCGCCAGTC 2508
Db 637 SerMetValLeuGlySerGlySerGlyValPheLysCysGlyIleAlaValAlaProVal 656
Qy 2509 ACTCTGTGATCTCTATGATACAGGATACACGGAACTGTTATAGGCT-----CACCT 2562
Db 657 SerArgTrpGluTyrTyrAspSerValTyrGluArgTyrMetGlyLeuProThrPro 676
Qy 2563 GACCAGAAATGAACAGGCTATTACTAGGATCTGTGCCCATGCAAGCAAGAAAGTTCCCC 2622
Db 677 GluAspAsnLeuAspHisTyrArgAsnSerThrValMetSerArgAlaGluAsnPhe--- 695
Qy 2623 TCTGAACCAATCGTTTACTCTTACATGTTTCTCGATGAGAAATGTCATTTGCA 2682
Db 696 ---LysGlnValGluTyrLeuLeuIleHisGlyThrAlaAspAspAsnValHisPheGln 714
Qy 2683 CATACCGATATATTACTGAGTGTTTTATGAGGCTGGAAAGCCATATGATTTACAGATC 2742
Db 715 GlnSerAlaGlnIleSerLysAlaLeuValAspValGlyValAspPheGlnAlaMetTrp 734
Qy 2743 TATCTCAGGAGACACAGCATAGAGTTCTCTGAAATCGGAGAACATTATGAACTGCAT 2802
Db 735 TyrThrAspGluAspHisGlyIleAlaSerSerThrAlaHisGlnHisIleTyrThrHis 754
Qy 2803 CTTTGTGCTACCTTCAAGAA 2823
Db 755 MetSerHisPheIleLysGln 761
RESULT 10
US-11-208-288-6
; Sequence 6, Application US/11208288
; Publication No. US20060051366A1
; GENERAL INFORMATION:
; APPLICANT: CHANG, Chiwen
; TITLE OF INVENTION: USE OF SOLUBLE CD26 AS INHIBITOR OF
; ANGIOGENESIS AND INFLAMMATION
; FILE REFERENCE: 39533-0001
; CURRENT APPLICATION NUMBER: US/11/208,288
```

```
; CURRENT FILING DATE: 2005-08-18
; PRIOR APPLICATION NUMBER: US 60/605,013
; PRIOR FILING DATE: 2004-08-26
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 760
; TYPE: PRT
; ORGANISM: Mus musculus
; US-11-208-288-6
Alignment Scores:
Pred. No.: 5,826-37 Length: 760
Score: 517.50 Matches: 203
Percent Similarity: 38.1% Conservative: 107
Best Local Similarity: 25.0% Mismatches: 306
Query Match: 9.3% Indels: 197
DB: 7 Gaps: 37
US-10-825-632-2 (1-3120) x US-11-208-288-6 (1-760)
Qy 604 ATGTATTCTCGAGAAAGAACTATTAAAGAAAAAGAAACGATTTGGAACAGTCGGAATT 663
Db 67 LeuTyrLysGlnGluAsnIleLeuLeuAsnAlaGluHisGlyAsnSerSerIle 86
Qy 664 -----GCTTCTTACGATTATCACCAGGAAGTGGAAACATTTCTG 702
Db 87 PheLeuGluAsnSerThrPheGluSerPheGlyTyrHisSerValSerProAspArgLeu 106
Qy 703 TTT-----CAAGCCGGTAGT 717
Db 107 PheValLeuLeuGluTyrAsnTyrValLysGlnTrpArgHisSerTyrThrAlaSerTyr 126
Qy 718 GGAATTTATCAGTAAAGATGGAGGGCCACAGGATTTACGCAACAACCTTTAAGGCC 777
Db 127 AsnIleTyrAspValAsnLys-----ArgGln 135
Qy 778 ATCTAGTGGAACTAGTTGTCCCAACATACGATGGATCCAAATTTATGCCCGCTGAT 837
Db 136 LeuIleThrGluGluValIleProAsn-----Asn 145
Qy 838 CCAGACTGGATT-----GCTTTATACATAGCAACGAT 870
Db 146 ThrGlnTrpIleThrTyrSerProGluGlyHisLysLeuAlaTyrValTrpLysAsnAsp 165
Qy 871 ATTGGATATCAACATCGTAACACAGAGAAGAGAGACTCACTTATGTGCACATGAG 930
Db 166 IleTyrValLysValGluProHisLeuProSerHisArgIleThr-----180
Qy 931 CTAGCCAACATCGAAGAGATGCCAGATCAGCTGAGTCGCTACTTTGTTCTCAAGAA 990
Db 181 ---SerThrGlyGluGluAsnValIleTyrAsnGlyIleThrAspTrpValTyrGluGlu 199
Qy 991 GAA---TTTGATAGATATTCTGGCTATTGGTGGTGTCACAAA-----1029
Db 200 GluValPheGlyAlaTyrSerAlaLeuTrpTrpSerProAsnAsnThrPheLeuAlaTyr 219
Qy 1030 GCTGAAACAACTCCCAAGTGGTGTAAATCTTGAATTTCTATATGAAGAAATGATGAA 1089
Db 220 AlaGlnPheAsnAspThrGlyValProLeuIleGluTyrSerPhe---TyrSerAspGlu 238
Qy 1090 TCTGAGGTGGAATATTATTCATGTTACATCCCTATGTTGGAACAAGAGGCGGAGATTC 1149
Db 239 Ser-----LeuGlnTyrProLysThrValTrp 247
Qy 1150 TTCCGGTTATCTAAACAGGTACACAAATCTTAAAGTCACTTTTAAGATGTCAGAAATA 1209
Db 248 IleProTyrProLysAlaGlyAlaValAsnProThrValLysPhePheIle-----Val 265
Qy 1210 ATGATTGATGCTGAGGAAGGATCATAGATGTCATAGATAAGGAACATAATCAACCTTTT 1269
Db 266 AsnIleAspSerLeuSerSerSerSerSerAlaAlaProIleGlnIleProAlaProAla 285
```

```
OY 1270 GAGATTCTATTGAGGAGTGTGAATATATTGCCAGAGCTGGATGGACTCTCTGAGGGA AAA 1329
   : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| |||
Db 286 SerValala---ArgGlyAspHisTyrLeuCysAspValValTTPAlaThrGluGluArg 304
OY 1330 TATGCT-----TGGTCCATCTACTAGATCGCTCCAG 1362
   : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| |||
Db 305 IleSerLeuGlnTrpLeuArgArgIleGlnAsnTyrSerValMetAlaIleCysAspTyr 324
OY 1363 ACTGCGCTACAGATAGTGTGATCTCACCCTGAAATATTATTATCCAGTAGAAGATGATTT 1422
   : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| |||
Db 325 AspLysIleAsnLeuThrTrpAsnCysProSer----- 335
OY 1423 ATGGAAGGAGAGACTCATTGAGTCAGTCGCTGATTCTGTGAGCCCACTAATTATCTAT 1482
   ||| : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| |||
Db 336 ---GluGlnGlnHisVal----- 340
OY 1483 GAAGAAACAACAGACATCTGGATAAATATCCATGACATCTTTTCATGTTTTTCCCAAGT 1542
   ||| : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| |||
Db 341 GluMetSerThrThrGlyTrpValGly-----ArgPheArgProAlaGluProHisPhe 358
OY 1543 CACGAAGAGAAATTGAGTTTATTATTTTGGCTCTGAATGCAAAACAGGTTTCCGTCATTTA 1602
   : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| |||
Db 359 ThrSerAspGlySerSerPheTyrLysIleIleSerAspLysAspGlyTyrLysHisIle 378
OY 1603 TACAAAATTACATCTATTTTAAAGGAACAAATATAACGATCCAGTGGTGGCTGCT 1662
   : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| |||
Db 379 CysHisPhe----- 381
OY 1663 GCTCAAGTGTATTCAG---TGCTCTATCAAGAGAGATAGCAATTACAGTGGTGA 1719
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 382 ---ProLysAspLysAspCysThrPhe-----IleThrLysGlyAla 395
OY 1720 TGGGAAGTTCTTGGCCGCATGATCTAATATCCAAATTGATGAAGTCAGAGCTGGTA 1779
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 396 TrpGluValIle-----SerIleGluAlaLeuThrSerAspTyrLeuTyr 410
OY 1780 TATTTTGAAGGCACC---AAAGACTCCCTTTAGAGCATCACCTCTAGTAGTCAGTTAC 1836
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 411 TyrIleSerAsnGlnTyrLysGluMetProGlyGlyArgAsnLeuTyrLysIle----- 428
OY 1837 GTAAATCTCGAGAGGTGACAGGCTGACTGACCGTGGCTACTCACATTTCTTGTCATC 1896
   : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| |||
Db 429 -----GlnLeuThrAsp-----HisThrAsnValLysCysLeu 439
OY 1897 AGT-----CAGCACTGTGACTTCTTTTATAAGTAAGTATAGTAACCAAG 1941
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 440 SerCysAspLeuAsnProGluArgCysGlnTyrTyrAlaValSerPheSerLysGluAla 459
OY 1942 AATCCACAC-----TGT-----GTGTCCTTTTACAAGCTATCAAGTCCT 1980
   : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| |||
Db 460 LysTyrTyrGlnLeuGlyCysTyrGlyProGlyLeuProLeuTyrThrLeuHisArgSer 479
OY 1981 GAAGATGACCACTTCCAAACAAGGAATTTTGGCCACCATTTTGGATTCAGCAGGT 2040
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 480 ThrAspHisLysGluLeuArgValLeuGlu-----AspAsnSerAla 493
OY 2041 -----CCTCTTCTGACTATCTCTCCAGAA-----ATTTTCTCTTTTGA 2082
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 494 LeuAspArgMetLeuGlnAspValGlnMetProSerLysLysLeuAspPheIleValLeu 513
OY 2083 AGTACTAGTGGATTTACATTTGATGGATGCTCTACAAGCCCTCATGATCTACAGCTGGA 2142
   : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| |||
Db 514 AsnGluThrArgPhe-----TrpTyrGlnMetIleLeuProProHis---PheAspLysSer 531
OY 2143 AAGAAATATCTACTGTGCTTCAATATATGTTGCTCTCAGGTGAGTGTGGTAAT 2202
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 532 LysLysTyrProLeuLeuLeuAspValTyrAlaGlyProCysSerGln----- 547
OY 2203 CGGTTTAAAGAGTCAAGTATTTCCGCTGGAAT-----ACCTAGCTCTCTAGGT 2253
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 548 -----LysAlaAspAlaSerPheArgLeuAsnTrpAlaThrTyrLeuAlaSerThrGlu 565
OY 2254 TATGTGGTTGTAGTG-----ATAGACAACAGGGGATCTCTGTCACCGAGGGCTTAATTTGAA 2310
```

```
Db 566 AsnIleIleValAlaSerPheAspGlyArgGlySerGlyTyrGlnGlyAspLysIleMet 585
OY 2311 GGCGCCTTTAAATATAAATGGGTCAAATAGAAATTCAGATCAGTGCAGGAGGAGCTCCAA 2370
   ||| : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| |||
Db 586 HisAlaIleAsnArgArgLeuGlyThrLeuGluValGluAspGlnIleGluAlaAlaArg 605
OY 2371 TATCTAGCTTCTCGATATGATTTTATTCATTCATGATCGTGTGGCATCCAGCGGTGGTCC 2430
   : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| |||
Db 606 GlnPheVal---LysMetGlyPheValAspSerLysArgValAlaIleTrpGlyTrpSer 624
OY 2431 TATGAGGATACCTCTCCCTGATGGCATTAATGCAGAGGTTCAGATATCTTCAGGTTGCT 2490
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 625 TyrGlyGlyTyrValThrSerMetValLeuGlySerGlySerGlyValPheLysCysGly 644
OY 2491 ATTCTGGGCCCCAGTCACTCTGTGGATCTCTCTATGATATACAGGATACACGAACGTTAT 2550
   : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| |||
Db 645 IleAlaValAlaProValSerArgTyrGluTyrTyrAspSerValTyrThrGluArgTyr 664
OY 2551 ATGGGT-----CACCTGACCAAGATGAACAGGCGCTATTACTTAGGATCTGTGCCCATG 2604
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 665 MetGlyLeuProIleProGluAspAsnLeuAspHisTyrArgAsnSerThrValMetSer 684
OY 2605 CAACACAGAAAGTTCCCTCTGAACCAATCGTTTACTGCTCTTACATGTTTCTCGAT 2664
   : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| |||
Db 685 ArgAlaGluHisPhe-----LysGlnValGluTyrLeuLeuIleHisGlyThrAlaHis 702
OY 2665 GAGATGTCATTTTGGCACAATACCATCTATATTACTGAGTGTTCAGTGGCTGGAAG 2724
   : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| |||
Db 703 AspAsnValHisPheGlnGlnSerAlaGlnIleSerLysLeuValAspAlaGlyVal 722
OY 2725 CCATATGATTTACAGATCTATCTCTCAGGAGACACAGCATAGAGTTCTCTGAATCGGGA 2784
   : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| |||
Db 723 AspPheGlnAlaMetTrpTyrThrAspGluAspHisGlyIleAlaSerSerThrAlaHis 742
OY 2785 GAACATATGAACTGCAATCTTTTTCACACTACCTTCAAGAA 2823
   : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| ||| : : : ||| |||
Db 743 GlnHisIleTyrSerHisMetSerHisPheLeuGlnGln 755

RESULT 11
US-11-186-284-55
; Sequence 55, Application US/11186284
; Publication No. US20050266493A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Berger, Allison
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Thibodeau, Stephen N.
; APPLICANT: Burgart, Lawrence J.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; FILE OF INVENTION: THERAPY OF COLON CANCER
; FILE REFERENCE: MPW01-0292P2RN
; CURRENT APPLICATION NUMBER: US/11/186,284
; CURRENT FILING DATE: 2005-07-21
; PRIOR APPLICATION NUMBER: US/10/301,822
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 60/339,971
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/361,978
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/381,988
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 55
; LENGTH: 760
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-11-186-284-55
```

Alignment Scores: 2.15e-31 Length: 760
 Pred. No.: 454.00 Matches: 169
 Score: 40.2% Conservativeness: 113
 Percent Similarity: 24.1% Mismatches: 269
 Best Local Similarity: 8.2% Indels: 150
 Query Match: 7 Gaps: 27
 DB:

US-10-825-632-2 (1-3120) x US-11-186-284-55 (1-760)

QY 847 ATTCTTTTATACATGACACGATATTTGGATATCTACATCGTAAACACAGAGAAGAAGG 906
 DB 162 LeuAlaTyRValTyRGlInAenAenIleTyRLeuLysGlnA-gProGlyAsPProPhe 181
 QY 907 AGACTCACTTATGTGCACATGAGCTAGCCACATGGAAGAGATGCCAGATCAGCTGGA 966
 DB 182 GlnIleThrPhe-----AsnGlyA-gGluAsnLysIlePheAsnGly 195
 QY 967 GTCGCTACCTTTGTTCTCCAAAGAAGATTT-----GATAGATATTCGGCTATTGGTG 1020
 DB 196 IleP-roAsPTrpValTyRGlInGluGluMetLeuProThrLysTyRAla---LeuTrpTrp 214
 QY 1021 TGTCCAAAGCTGAAACAACTCCAGTGGTGTGTAATAATCTTAGAATCTTATAGAA 1080
 DB 215 SerPro-----AsnGlyLysPheLeu-----AlaTyRAlaGlu 225
 QY 1081 AATGATGAATCTGAGGTGGAATATTATTCATGTTACATCCCTATGTTGGAACA---AGG 1137
 DB 226 PheAsnAspLysAspIleProValIleAlaTyRSerTyR-TyR-GlyAspGluGlnTyRPro 245
 QY 1138 AGGCGAGATTCATCCGTTATCTTAAACAGGTACAGCAATCCCTAAAGTCACTTTTAA 1197
 DB 246 ArgThrIleAenIleProTyRProLysAlaGlyAlaLysAsnProValValArgIle--- 264
 QY 1198 ATGTCAGAAATATGATGCTGGAAGGAGATCATAGATCTCATAGATGAAGAACTA 1257
 DB 265 -----PheIleAspThrThrTyRProAlaTyR 274
 QY 1258 ATCAACCTTTTGAGATT-----CTATTGAGGAGTTGAAATATTATTC 1302
 DB 275 ValGlyProGlnGluValProValProAlaMetIleAlaSerSerAspTyR-TyR-PheSer 294
 QY 1303 AGAGCTGGATGACTCCTGAGGGAATAATATGCTTGGTCCATCTCATAGATCGCTCCCG 1362
 DB 295 TrpLeuThrTrpValThrAspGluArgValCysLeuGlnTrpLeu----- 309
 QY 1363 ACTCGCTACATAGTGTGATCTCACCTGAATATTATTATCCAGTGAAGATGATGT 1422
 DB 310 LysArgValGlnAsnValSerValLeuSerIleCysAspPheArgGluAspTrpGlnThr 329
 QY 1423 ATGGAAGGCAGAGACTCATTCAGTCAGTCAGTCGCTGATTCCTGTGACG----- 1467
 DB 330 TrpAspCysProLysThrGlnGluHisIleGluGluSerArgThrGlyTrpAlaGlyGly 349
 QY 1468 -----CCACTAATTATCTATGAAGAAACACAGACATCTCGATAAATATC 1512
 DB 350 PhePheValSerArgProValPheSerTyRAspAla-----IleSerTyR 364
 QY 1513 CATGACATCTTTCATGCTTTTCCCAAGTCACGAGAGAAATTGAGTTATTATTGTC 1572
 DB 365 TyRlysIlePhe----- 368
 QY 1573 TCTGAATGCAAAACAGGTTTCGTCATTTATACAAATTTACATCTATTTTAAAGAAAGC 1632
 DB 369 ---SerAspLysAspGlyTyRlysHisIleHisTyRile----- 380
 QY 1633 AAATATAACGATCCAGTGGTGGCTGCTCCAGTGCCTCCAGTGCCTCCAGTGCCTCCAGT 1692
 DB 381 -----LysAspThrValGlu 385
 QY 1693 GAGGAGATAGCAATTACCACTGGTGAATGGGAAGTTCTTGGCCGCGCATGATCTAATATC 1752

DB 386 AsnAlaIleGlnIleThrSerGlyLysTrpGluAlaIle-----AsnIle 400
 QY 1753 CAAGTTGATGAAGTCAAGAGCTGGTATATTTTGAAGGCACCAAGACTCCCTTTAGAG 1812
 DB 401 -----PheArgValThrGlnAspSerLeuPheTyR 410
 QY 1813 CATCACCTGTACGTAGTCACTAGTAAATCTCGAGAGGTGACAGCTGACGACCGT 1872
 DB 411 SerSerAsnGluPheGluGluTyRProGlyA-gArgAsnIleTyRArgIleSerIleGly 430
 QY 1873 GGCTACTCACATTTCT---TGCTGCATCAGTCAGCAC-----TGTCACCTC 1914
 DB 431 SerTyRProSerLysLysCysValThrCysHisLeuA-gLysGluA-gCysGlnTyR 450
 QY 1915 TTTATAAGTAAATATAGTAAACAGAGAATCCACATCTGTGTGCTCCTTTAC----- 1965
 DB 451 TyRThrAlaSerPheSerAspTyRAlaLysTyRTrpAlaLeuValCysTyR-GlyProGly 470
 QY 1966 ---AAGCTATCAAGTCTGAGAGTACCCAACTTCGAAACAAAGAAATTTTGGGCCACC 2022
 DB 471 IleProIleSerThrLeuHisAspGlyA-gThrAspGlnGlu-----IleLys 486
 QY 2023 ATTTTGGATTTCAGCAGGTCTCTCTCTGAC-----TATACTCTCTCCAGAA 2067
 DB 487 IleLeuGluGluAsnLysGluLeuGluAenAlaLeuLysAenIleGlnLeuProLysGlu 506
 QY 2068 ATTTCTCTTTTGAAGTACTACTGGATTACATTTGATGATGCTCTACAAGCTCAT 2127
 DB 507 GluIleLysLysLeuGluValAspGluIleThrLeuTrpTyR-LysMetIleLeuProPro 526
 QY 2128 GATCTACAGCTGGAAGAAATATCTCTACTGCTGCTTTCATATATGTTGCTCCTCAGGTG 2187
 DB 527 GlnPheAspArgSerLysTyRProLeuLeuIleGlnValTyR-GlyGlyProCysSer 546
 QY 2188 CAGTTGGTGAATATCGGTTTAAAGGAGTCAAGTATTTCCGTTGATACCTCAGCTCT 2247
 DB 547 GlnSerValArgSer-----ValPheAlaValAsnTrpIleSerTyR 560
 QY 2248 CTA-----GGTTATGTGTTGTAGTAGACACACAGGGGATCCTGTCCCGA 2295
 DB 561 LeuAlaSerLysGluGlyMetValIleAlaLeuValAspGlyArgGlyThrAlaPheGln 580
 QY 2296 GGGCTTAAATTTGAAGCGCTTTTAAATATAAATGGGTCAATAGAAATGACGATCAG 2355
 DB 581 GlyAspLysLeuLeuTyRAlaValTyRArgLysLeuGlyValTyRGlValGluAspGln 600
 QY 2356 GTGGAAGGACTCCAATATCTAGCTTCTCGATATGATTTCACTTCACTTAGATCGTGGGC 2415
 DB 601 IleThrAlaValArgLysPheIle-----GluMetGlyPheIleAspGlyLysArgIleAla 619
 QY 2416 ATCCAGCTGTCTCTATGAGGAGTACCTCTCTGATGGCATTTAATGACAGGTGATGAT 2475
 DB 620 IleTrpGlyTrpSerTyR-GlyGlyTyRValSerSerLeuAlaLeuAlaSerGlyThrGly 639
 QY 2476 ATCTTCAGGTTGCTTATGCTGGGCGCCAGTCACTCTGTGGATCTTCTATGATACAGGA 2535
 DB 640 LeuPheLysCysGlyIleAlaValAlaProValSerSerTrpGluTyRTrpAlaSerVal 659
 QY 2536 TACACGGAACGTTATATGGTCACCT-----GACCAGATGAACAGGCTATTACTTA 2589
 DB 660 TyRThrGluArgPheMetGlyLeuProThrLysAspAspAsnLeuGluHisTyR-LysAsn 679
 QY 2590 GGAATCTGGCCCATGACAGCAAAAGTCTCCCTCTGAACCAATCGTTTACTCTCTTA 2649
 DB 680 SerThrValMetAlaArgAlaGluTyR-PheArgAsnValAsp-----TyRLeuLeuIle 697
 QY 2650 CATGGTTCTCGGATGAGAAATGTCATTTTTCACATACCATCATATATGAGTTTTTTA 2709
 DB 698 HisGlyThrAlaAspAsnValHisPheGlnAsnSerAlaGlnIleAlaLysAlaLeu 717
 QY 2710 GTGAGGCTCGAAAGCCATATATGATTTACAGATCTATCTCTCAGGAGACACAGCATAGA 2769
 DB 718 ValAsnAlaGlnValAspPheGlnAlaMetTrpTyRSerAspGlnAsnHisGlyLeu--- 736

QY 2770 GTTCTGAATCGGGA-----GAACATTATGAATGCACTCTTTGGCACTACCTTCAA 2820
Db 737 -----SerGlyLeuSerThrAsnHisLeuTyrThrHisMetThrHisPheLeuLys 753
QY 2821 GAA 2823
Db 754 Gln 754

RESULT 12

US-11-176-951-11
; Sequence 11, Application US/11176951
; Publication No. US20060024313A1
; GENERAL INFORMATION:
; APPLICANT: CHEN, XIN
; TITLE OF INVENTION: AGENTS THAT DISRUPT DIMER FORMATION IN DPP-IV FAMILY OF
; TITLE OF INVENTION: AGENTS THAT DISRUPT DIMER FORMATION IN DPP-IV FAMILY OF
; FILE REFERENCE: 08842.0019
; CURRENT APPLICATION NUMBER: US/11/176,951
; CURRENT FILING DATE: 2005-07-06
; PRIOR APPLICATION NUMBER: 60/586,095
; PRIOR FILING DATE: 2004-07-06
; PRIOR APPLICATION NUMBER: 60/585,952
; PRIOR FILING DATE: 2004-07-06
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patent In Ver. 3.3
; SEQ ID NO 11
; LENGTH: 99
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-176-951-11

Alignment Scores:
Pred. No.: 2,03e-26 Length: 99
Score: 394.00 Matches: 72
Percent Similarity: 82.8% Conservative: 10
Best Local Similarity: 72.7% Mismatches: 17
Query Match: 7.1% Indels: 0
Gaps: 0

US-10-825-632-2 (1-3120) x US-11-176-951-11 (1-99)

QY 2533 GGATACCGGAAGCTTATATGGTCACTGACCAAGATGAACAGGCTATTACTTAGGA 2592
Db 1 GlyTyrThrGluArgTyrMetAspValProGluAsnAsnGlnHisGlyTyrGluAlaGly 20
QY 2593 TCTGTGGCCATGCAACAGAAAGTTCCTCTGAACCAATCGTTTACTCTCTTACAT 2652
Db 21 SerValAlaLeuHisValGluLysLeuProAsnGluProAsnArgLeuLeuLeuHis 40
QY 2653 GGTTCCTCGATCAGATGCTCCATTTTGCACATACCATATATATTACTGAGTTTATTAGTG 2712
Db 41 GlyPheLeuAspGluAsnValHisPheHisThrAsnPheLeuValSerGlnLeuLeu 60
QY 2713 AGGCTGGAAGCCATATGATTTTACAGATCTATCTCAGAGACACAGACATAGAGTT 2772
Db 61 ArgAlaGlyLysProTyrGlnLeuGlnIleTyrProAsnGluArgHisSerIleArgCys 80
QY 2773 CCTGAATCGGAGAACATTATGCACTGTCATCTTTTGGCACTACCTTCAAGAAACCTT 2829
Db 81 ProGluSerGlyGluHisTyrGluValThrLeuLeuHisPheLeuGlnGluTyrLeu 99

RESULT 13

US-11-079-463-7504
; Sequence 7504, Application US/11079463
; Publication No. US20060073161A1
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTERIOIDES FRA
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: PATH00-03DIV2
; CURRENT APPLICATION NUMBER: US/11/079,463

; CURRENT FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/128,705
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: US 09/540,209
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 7504
; LENGTH: 624
; TYPE: PRT
; ORGANISM: B.fragilis
US-11-079-463-7504
Alignment Scores:
Pred. No.: 5,62e-14 Length: 624
Score: 255.00 Matches: 161
Percent Similarity: 34.1% Conservative: 103
Best Local Similarity: 20.8% Mismatches: 244
Query Match: 4.6% Indels: 267
Gaps: 36
DB:
US-10-825-632-2 (1-3120) x US-11-079-463-7504 (1-624)
QY 681 CCAAGGAAGTGGAAACATTTCTGTTTCAAGCCGGTAGTGGAAATTTATCAGTAAAGATGG 740
Db 1 ProArgTyrTrpAsp-----AsnHisSerLeuLysArg--- 11
QY 741 AGGCCACAAGGATTTACGCAACACCTTTAAGGCCCAATCTAGTGGAAATCTAGTTGCC 800
Db 12 -----SerArgThrTyrCysGluLeuThrGluLeuLysThrGlyLysValLeuLeuThr 29
QY 801 CAACATACGATGATCCAAATATTATGCCCCTGATCCAGACTGATGCTGTTTATACA 860
Db 30 AsnLeuArgAspGlyMetArgTrpMetProLysSerAsnLysLeu-----TyrTyrThr 47
QY 861 TAGC-----AACGATATTGGATATCTAACATCTGAACCCAGACAGAGAAG 905
Db 48 -ValValAlaProGluGlyAsnAspValIleThrLeuAspProValThrLeuLysGluGly 67
QY 906 GAGACTCACCTATGTGCACAAATGAGCTAGCCAACTGGAAGAGATGCCAGATCAGCTGG 965
Db 67 uValLeu-----LeuArgGlyIleProGlu----- 75
QY 966 AGTCGCTACCTTTGTTCTCCAGAGAATTTGATAGATATCTGGCTATTGGTGGTGGTCC 1025
Db 76 -----GlnGlyPheSerTrpSerPr 82
QY 1026 AAAAGCTGAACAACACTCCAGTGGTGGTAAATTTCTTAGAATTTCTATAT-----GAAGA 1079
Db 82 AsnGluAsp-----PheLeuIleTyrTyrProArgGluGly 94
QY 1080 AAATGATGAATCTGAGTGGAAATATTATCATGTTATACATCCCT-----ATGTTGGAAC 1133
Db 94 uGlyValLysAspGluGlyProLeuLysArgIleValSerProAlaAspArgIleProAs 114
QY 1134 AAGGAGGCGAGATTCATTC-----CGTTATCTTAAACAGGTACAGCAATCTTAAAGT 1187
Db 114 nThrArgGlyArgSerPheLeuAlaArgTyrAspIleAlaSerGlyThrSerGluArgLe 134
QY 1188 CACTTTT-----AAGATGTCAGAAATATGATGATGCTGGAAGGAGGAT 1232
Db 134 uThrTyrGlyAsnHisSerThrTyrMetGlnAspIleSerProAspGlyLysTyrLeuLe 154
QY 1233 CATAGATGTCATAGTAAGGAACCTAATCAA---CCTTTTGTAGATT----- 1275
Db 154 uTyrSerSerSerLysGluAsnIleThrGlnArgProPheSerLeuSerSerLeuPheGly 174
QY 1276 -----CTATTGAAGGAGTGAATATATTATTCGCCAG 1304
Db 174 nValAsnLeuGluThrLeuAlaValAspThrLeuPhePheGluAspArgPheLeuGlyGly 194
QY 1305 AGCTGGATGAGACTCTGAGGAGGAAAAATATGCTTGGTCCATCTTACTAGATCGCTCCACAG 1364
Db 194 yAlaSerTyr-SerProAspGlyLys----- 202


```
Db      83  AspArgGluGlyAspAlaGlnLeuTyrIleMetSerThrGluGlyGly-----Glu 100
Qy      1132  ACAAGAGGGCAGATTTCATCCGTTATTCCTAAACAGAGTACAGCAAAATCTTAAGTCACT 1191
Db      101  AlaArgLysLeuThrAspIleProTyr-----GlyValSerLysProLeuTrpSer 117
Qy      1192  TTTAAGATGTCAGAAATAATGATTGAT-----GCTGAAGGAAGATCATAGATGTC 1242
Db      118  ProAspGlyGluSerIleLeuValThrIleSerLeuGlyGluGlyGluSerIleAspAsp 137
Qy      1243  ATAGATAAG-----GAACTAATTCAACCTTTTGAGATTCTATTGTAAGGAGTGGAA 1293
Db      138  ArgGlnLysThrGluGlnAspSerTyrGluProValGluVal-----GlnGlyLeuSer 155
Qy      1294  TATATTGCCAGAGCTGGATGGACTCCTGAGGGAAAAATATGCTGTGTCCTACTAGAT 1353
Db      156  TyrLysArgAspGlyLysGlyLeuThrArgGlyAlaTyrAla----- 169
Qy      1354  CGCTCCAGACTCGCCTACAGATAGTGTGTGATCTCACCTGATTCATTTATCCAGTAGAA 1413
Db      170  -----GlnLeuValLeuValSerValLys----- 177
Qy      1414  GATGATGTTATGGAAGGAGAGACTCATGATGATGAGTGCCTGATCTGTGACGCCACTA 1473
Db      178  -----SerGlyGluMetLysGluLeuThrSerHisLysAlaAspHisGlyAspProAla 195
Qy      1474  -----ATTATCTATGAGAAACA-----ACAGACATCTGG 1503
Db      196  PheSerProAspGlyLysTrpLeuValPheSerAlaAsnLeuLeuThrGluThrAspAspAla 215
Qy      1504  ATAAATATCCATGATCTTTCATGTTTCCCAAGTCACGAGAGGAAATGAGTTT 1563
Db      216  SerLysProHisAspValTyr-----IleMetSer 225
Qy      1564  ATTTTTCCTCTCAATGCAAA-----ACAGGTTTCCGTCATTATATACAAATACATCT 1617
Db      226  LeuGluSerGlyAspLeuLysGlnValThrProHisArgGlySerPheGlySerSerSer 245
Qy      1618  ATTTTAAAGGAACCAATTAACGATCCAGTGGGCTGCTGCTGCCAAGTGATTTC 1677
Db      246  PheSerProAspGlyArgTyrLeuAlaLeuLeuGlyAsn-----GluLysGluTyr 262
Qy      1678  AAGTGTCTTATCAAGAGGAGATAGCAATTACAGTGGTGATGG----- 1722
Db      263  Lys-----AsnAlaThrLeuSerLysAlaTrpLeuTyrAspIleGlu 276
Qy      1723  -----GAAGTTCTTGCCGGCAT----- 1740
Db      277  GlnGlyArgLeuThrCysLeuThrGluMetLeuAspValHisLeuAlaAspAlaLeuIle 296
Qy      1741  GGATCTAATATCCAGTTGATGAAGTCAGAGGCTGGTATATTTTGAAGGCACCAAGAC 1800
Db      297  GlyAspSerLeuIleGlyGlyAlaGluGlnArgProIleTrp-----ThrLysAsp 313
Qy      1801  TCCCTTTAGAGCATCACCTGTAGTACGTAGTCAGTTAGTAAATCCTGGAGAGGTGACAAG 1860
Db      314  Ser-----GlnGlyPheTyrValIleGly----- 321
Qy      1861  CTGACTGACCGTGGC----- 1875
Db      322  ---ThrAspGlnGlySerThrGlyIleTyrTyrIleSerIleGluGlyLeuValTyrPro 340
Qy      1876  -----TACTCACATTTCTGTGTCATCAGTCAGCAGCTGTGACTTCTTT 1917
Db      341  IleArgLeuGluLysGluTyrIleAsnSerPheSerLeuSerProAspGluGlnHisPhe 360
Qy      1918  ATAGTAGATATAGTACCAAGAGATCCACACTGTGTCTCCTTTACAGCTATCAAGT 1977
Db      361  IleAlaSerValThrLysProAspArgProSerGluLeuTyrSerIleProLeuGlyGln 380
Qy      1978  CTTGAAGATGACCCCACTTCGAAACAAAGGAATTTTGGCCACCATTTTGGATTCAGCA 2037
```

```
Db      381  GluGluLysGlnLeuThrGlyAlaAsnAspLysPhe----- 392
Qy      2038  GGTCTCTCTCCTGACTATACT-----CCTCCAGAAATTTTCTCTTTTCAAGTACTACT 2091
Db      393  -----ValArgGluHisThrIleSerIleProGluGluIleGlnTyrAlaThrGluAsp 410
Qy      2092  GGATTTTACATTTGATGGGATGCTCTACAAGCCTCATGATCTACAGCCTGGAAAGAAATAT 2151
Db      411  GlyValMetValAsnGlyTrpLeuMetArgProAlaGlnMetGluGlyGluThrThrTyr 430
Qy      2152  CCTACTGTGCTGTTCATATATGATGCTCTCAGGTGCGATTCGTGGAATAATCGGTTTAAA 2211
Db      431  ProLeuIleLeuAsnIleHisGlyGlyProHisMetMetTyr----- 444
Qy      2212  GGAGTCAAGTATTTC---CGCTTGATATACCTAGCTCTCTAGGTATATGTGGTCTAGTG 2268
Db      445  GlyHisThrTyrPheHisGluPheGlnValLeuAlaAlaLysGlyTyrAlaValValTyr 464
Qy      2269  ATAGACAACAGGGGATCCTGTCAACCGAGGGCTTAATTTGAAGGCGCTTTTAAATATAAA 2328
Db      465  IleAsnProArgGlySerHisGlyTyrGlyGlnGluPheValAsnAlaValArgGlyAsp 484
Qy      2329  ATGGGTCAAAATAGAAATTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATAT 2388
Db      485  TyrGlyGlyLysAspTyrAspValMetGlnAlaValAspGluAlaIleLysArgAsp 504
Qy      2389  GATTTTCACTTACATCTAGATCCTGTGGGCATCCAGCGTGTCTATGGAGGATACCTCTCC 2448
Db      505  ProHisIleAspProLysArgLeuGlyValThrGlyGlySerTyrGlyGlyPheMetThr 524
Qy      2449  CTGATGGCATTAATGCAAGGTCAGATATCTTCAGGTTGCTATTGCTGGGCCCCCAGTC 2508
Db      525  AsnTrp---IleValGlyGlnThrAsnArgPheLysAlaAlaValThrGlnArgSerIle 543
Qy      2509  ACTGTGTGATCTCTAT-----GATACAGGATAC-----ACGGAACGTTAT 2550
Db      544  SerAsnTrpIleSerPheHisGlyValSerAspIleGlyTyrPhePheThrAspTrpGln 563
Qy      2551  ATGGGTCAACCT-----GACCAGAATGAACAGGCTAT 2583
Db      564  LeuGluHisAspMetPheGluAspThrGluLysLeuTrpAspArgSerProLeuLysTyr 583
Qy      2584  TACTTAGGATCTGTGGCGCATGCAAGCAGAAAGTCCCTCTGAAACCAATCGTTTACTG 2643
Db      584  -----AlaAlaAsnValGlu-----ThrProLeuLeu 592
Qy      2644  CTCTTACATGTTTCTCTGGATGAGAATGTCCATTTTGCATACCCAGTATATATTCTGAGT 2703
Db      593  IleLeuHisGlyGluArgAspAspArgCysProIleGluGlnAlaGluGlnLeuPheIle 612
Qy      2704  TTTTGTAGTGGGCTGGAAAGCCATATGATTTACAGATCTATCTCTCAGAGAGACACAGC 2763
Db      613  AlaLeuLysLysMetGlyLysGluThrLysLeuValArgPheProAsnAlaSerHisAsn 632
Qy      2764  ATA 2766
Db      633  Leu 633
```

RESULT 15

US-11-079-463-9281

; Sequence 9281, Application US/11079463

; Publication No. US20060073161A1

; GENERAL INFORMATION:

; APPLICANT: Gary L. Breton

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTERIOIDES FR

; FILE REFERENCE: PATH00-03DIV2

; CURRENT APPLICATION NUMBER: US/11/079,463

; CURRENT FILING DATE: 2005-03-14

; PRIOR APPLICATION NUMBER: US 60/128,705

; PRIOR FILING DATE: 1999-04-09

; PRIOR APPLICATION NUMBER: US 09/540,209

; PRIOR FILING DATE: 2000-04-04

; NUMBER OF SEQ ID NOS: 10444		
; SEQ ID NO 9281		
; LENGTH: 737		
; TYPE: PRT		
; ORGANISM: B.fragilis		
US-11-079-463-9281		
Alignment Scores:		
Pred. No.:	3,56e-09	Length: 737
Score:	200.50	Matches: 143
Percent Similarity:	32.9%	Conservative: 107
Best local Similarity:	18.8%	Mismatches: 276
Query Match:	3.6%	Indels: 233
DB:		Gaps: 31
US-10-825-632-2 (1-3120) x US-11-079-463-9281 (1-737)		
QY	862	AGCAACGATTTGGATATCTAATCATCGTAAACAGAGAAAGAGAGACTCACT-----915
DB	117	AsnArgGluValPheValMetAsnAlaAaspGlyThrAspAsnArgGlnIleThrHisThr136
QY	916	---TATGTGCAATGAG-----930
DB	137	ProTyrGlnGluAsnGluValThrTrpAlaAaspGlySerLysLeuLeuPheLeuSer156
QY	931	-----CTAGCCAAACATCGAAGAAAGATGCCAGATCAGCTGGAGTC969
DB	157	AsnAspAsnGlySerSerGlnLeuTyrGluMetAsnProAspGlySerGly-----173
QY	970	GCTACCTTTGTCTCCAAAGAAAGAAATTTGATAGATAT-----TCTGGCTATTGG1017
DB	174	-----ArgLysGlnIleSerLysTyrAspGlyAspIleGluGlyTyr---187
QY	1018	TGGTCTCCAAAAGCTGAACAACTCCAGTGGTGGTAAATCTTAGAATCTATATGAA1077
DB	188	-----SerIleSerProAspGlyLysLysIleLeuPheIle-----199
QY	1078	GAAATGATGATCTGAGTGGAAATATTATCATGTTACATCCCTAGTGTGAAACAAAGG1137
DB	200	-----AlaGlnValLysThrVal205
QY	1138	AGGGCAGATTCTTCCTTATCCGTTATCTAAACACAGGTACAGCA-----1176
DB	206	LysSerThrAlaAspLysTyrProAspLeuAspLysAlaThrGlyIleIleIleThrAsp225
QY	1177	-----AATCCTAAAGTCACCTTTTAAGATGTCAGAAATAATGATTGAT1218
DB	226	LeuMetTyrLysHisTrpAspGluTrpValThrAlaProHisProPheIleAlaSer245
QY	1219	GCTGAAGGAAGG-----ATCATAGATGCTATAGATAAGGAAGTAATTCNACT1266
DB	246	PheAspGlyLysSerIleSerAsnIleIleAspValLeuGluGlyGlu-----Pro262
QY	1267	TTTGAGATTCTATTGAA-----GGAGTTGAATATATTGCCAGAGCTGATGG---1314
DB	263	TyrGluSerProMetLysProTrpGlyIleGluGlnLeuAla-----TrpAsn279
QY	1315	ACTCTGTAGGGGAAATATGCTTGTGTCCTACTAGATCGCTCCAGACTCGCTACAG1374
DB	280	ThrThrSerAspLysValAlaTyrThrCys-----ArgLysLysThrGlyLeuAla296
QY	1375	ATAGTGTGATCTCACCTGAATATTATTATCCAGTAGAAGATGATGTTATGGAAGGCGAG1434
DB	297	TyrAlaIle-----299
QY	1435	AGACTCATTTGAGTCAGTCGATCTGTGACGCCACTAATTATCTATGAAGAAACAA1494
DB	300	-----SerThrAsnSerAspIleTyrValTyrAspLeuAsnThr312
QY	1495	GACATCTGATTAATATCCATGAC-----ATCTTTGATGTTTTCCTCCAA-----1539
DB	313	LysLysThrValAsnIleThrGluGlyMetMetGlyTyrAspThrAsnProGlnTyrSer332
QY	1540	-----AGTCACGAAGAGGAATTT1557
DB	333	ProAspGlyLysSerIleAlaTrpGlnSerMetGluAtrgAspGlyTyrGluAlaaspGln352
QY	1558	GAGTTTATTTTTCCTCTGNAATGCAAAACAGAGTTTTCCTCATTTATACAAA-----ATT1611
DB	353	AsnArgLeuPheValMetAsnLeuGluThrGlyGluLysArgPheValSerLysAlaPhe372
QY	1612	ACATCTATTAAAGAAAGCAAAATATAACGATCCAGTGGTGGGCTGCTCCAAGT1671
DB	373	AspSerAsnValAspAlaPheValTrpSerArgAspAla-----GCTGAATGG385
QY	1672	GATTTCAAGTGTCTCTATCAAAAGAGGAGATAGCAATTTACCAGT-----1722
DB	386	-----LysThrIleTyrPheThrGlyValTrpHisGlyGluThr398
QY	1723	GAAGTTCCTTGGCCGGCATGATCTAATATCAACAGTTGATGAAGTCAGAGGCTGGTATAT1782
DB	399	GlnIleTyrSerLeuAspLeuSerAsnAspSerValArgProValThrSerGlyMetTyr418
QY	1783	---TTTGAAGGCACC-----AAAGACTCCCT1806
DB	419	AspTyrGluGlyValAlaLeuPheGlyAspLysLeuIleAlaLysArgHisSerMetSer438
QY	1807	TTAGAGCATCACCTGTAGTAGTCAGTTACGTAAATCCTGGAGAGGTGACAAGGCTGACT1866
DB	439	MetGlyAspGluIleTyrAlaIleAlaLeu-----AspGlyGlnThrThrGlnLeuThr456
QY	1867	GACCGTGGCTACTCACATTCTTGTGTCATCAGTCAGCACTGTGACTTCTTTATAAGTAAG1926
DB	456	-----456
QY	1927	TATAGTAACGAGAGAAATCCACACTGTGTCCTTTTACAAGCTATCAAGTCCTGAAGAT1986
DB	457	-----GlnGluAsnLysGlnIleTyrAspGlnIleGluMetGly-----469
QY	1987	GACCCAACTTGCAAAACAAAGGAATTTTGGCCACCATTTTGGATTCAGCAGGCTCCTCT2046
DB	470	-----LysValGluGlyArgTrpMetLysThrThrAspSerLysGluMetLeu485
QY	2047	CCTGACTACTCTCCAGAAATTTCTCTTTTGAAGTACTACTGGATTACATTGAT2106
DB	486	Thr-----TrpValIleTyr490
QY	2107	GGGATGCTCTACAGCCCTCATGATCTACAGCTCGAAGAAATATCTACTGTGCTGTTTC2166
DB	491	-----ProGlnPheAspProAsnLysLysTyrProThrLeuLeuPhe505
QY	2167	ATATATGTGTGTCCTCAGTGCAGTTGGTGAATTAATCGTTTAAAGAGTCAAGTATTTC2226
DB	506	CysGluGlyGlyProGlnSerProVal-----SerGlnPheTrpSerTyrArgTrp---522
QY	2227	CGCTTGAATATACCTAGCCCTCTCTAGTTATGTGGTTGTGTAGTAGATAGACAGGGGATCC2286
DB	523	AsnMetGlnIleMetAlaAlaAsnGlyTyrIleValValAlaProAsnArgArgGlyLeu542
QY	2287	TGTCACCGAGGCTTAAATTTGAGGCGCCTTTAAATATAAAATGGGTCAATAGAAATTT2346
DB	543	ProGlyPheGlyLeuGluTrpAsnGluAlaIleSerGlyAspTyrGlyGlnCysMet562
QY	2347	GACATCAGGTGGAAGGACTCCAAATATCTAGCTTCTCGATATGATTTTCACTGACTAGAT2406
DB	563	LysAspTyrPheThrAlaIleAspGluMetAla---LysGluProPheValAspSerAsp581
QY	2407	CGTGTGGGCTCCACGGTGGTCTTATGGAGATAC-----2442
DB	582	ArgLeuGlyCysValGlyAlaSerPheGlyGlyPheSerValTyrTrpLeuAlaGlyHis601
QY	2443	-----CTCTCCCTGATGGCATTAATGCGAGGTCAGATATCTTCACGGTTGCTTGTCT2496
DB	602	HisAspLysArgPheLysAlaPheIleAlaHisAspGlyIlePheAsnMetGluMetGln621
QY	2497	GGGGCCCCAGTCACTCTGTGGATCTTCTATGATACAGGATACACCGAAGCTTATATGGGT2556

```
Db      622 TyrLeuGluThrGluGluLysTrpPheAlaAsnTrpAspMetGlyGlyAlaTyTrpGlu 641
Qy      2557 CACCCTGACCAAGATGAACAGGGCTATTACTTAGGATCTGTGGCCATGCCAAGCAGAAAAG 2616
Db      642 LysGlnAsnProThrAlaGlnArgThrPheAlaAsnSerProHisLeuPheValGluLys 661
Qy      2617 TTCCCCTCTGAACCAATCGTTTACTGCTCTTACATGGTTTCTGATGGTTCCTGGATGAGAATGCCAT 2676
Db      662 TrpAspThrPro-----IleLeuCysIleHisGlyGluLysAspTyrArgIleLeu 678
Qy      2677 TTTCACATACCAGTATATTACTGAGTTTTTTTAGTCAGGGCTGGAAAGCCCATATGATTTA 2736
Db      679 AlaAsnGlnAlaMetAlaAlaPheAspAlaAlaValMetArgGlyValProAlaGluLeu 698
Qy      2737 CAGATCTATCTCAGGAGACACACAGCATAAGAGTTCTGTAATCGGGAGAACATTATCAA 2796
Db      699 LeuIleTyrProAspGluAsnHisTrpValLeuLysProGlnAsnGlyValLeuTrpGln 718
Qy      2797 CTGCATCTTTTGCACTACCTTCAAGAAAACCTTGGATCACGTTATTTGCTGCTCTAAA 2853
Db      719 ArgThrPhePheGluTrpLeuAspGlnTrpLeuLysProAsnGluThrAlaGlnLys 737
```

Search completed: April 14, 2006, 12:43:20
Job time : 109.5 secs

THIS PAGE BLANK (USPTO)

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☒ **BLACK BORDERS**
- ☒ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.

THIS PAGE BLANK (USPTO)